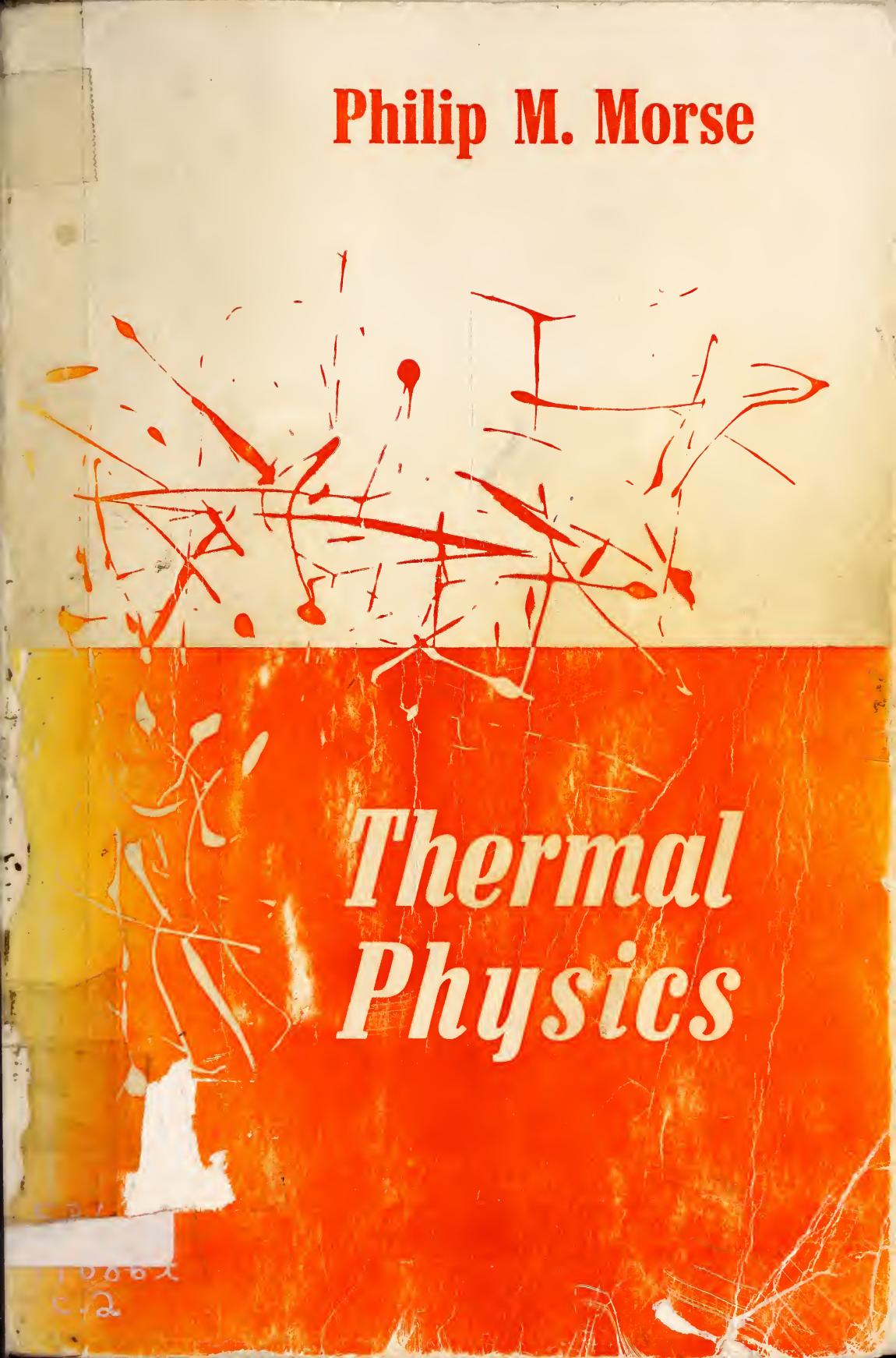


Philip M. Morse



*Thermal  
Physics*

100001  
c.2

UNIVERSITY  
OF FLORIDA  
LIBRARIES



PHYSICS  
SEMINAR

ENGINEERING AND PHYSICS

LIBRARY

# *Thermal Physics*



# ***THERMAL PHYSICS***

*A Preliminary Edition*

**Philip M. Morse**  
*Professor of Physics*  
*Massachusetts Institute of Technology*

**W. A. Benjamin, Inc.**  
New York                    1962

536,7  
M886t  
c. 2

ENGINEERING

& PHYSICS

LIBRARY

THERMAL PHYSICS  
A Preliminary Edition

Copyright © 1961 by W. A. Benjamin, Inc.  
All rights reserved.

Library of Congress Catalog Card Number: 61-18590  
Manufactured in the United States of America

*The manuscript was received  
on August 25, 1961, and published  
on January 25, 1962*

W. A. BENJAMIN, INC.  
2465 Broadway, New York 25, New York

# Preface

This book represents a stage in the process of revising the course in heat and thermodynamics presented to physics undergraduates at the Massachusetts Institute of Technology. It is, in its small way, a part of the revision of undergraduate curriculum in physics which has recently been going on at the Institute and at other centers of physics in this country. Such major revisions must be made at least once a generation if we are to compress the expanding field of physics into a form that can be assimilated by the physics student by the time he completes his graduate education.

Such a downward compaction of new physics, introducing it earlier in the education of a physicist, is not a task that can be undertaken one subject at a time. For example, the basic techniques of classical dynamics and electromagnetic theory could not now be taught effectively to juniors or seniors unless they were already fluently conversant with the differential calculus. This particular revision was accomplished, in this country, about a generation ago; prior to that time graduate students had to unlearn the geometric tricks they had learned as undergraduates, before they could assimilate the methods of Lagrange and Hamilton.

Now it is necessary to bring the concepts of the quantum theory into the undergraduate curriculum, so the graduate student does not have to start over again when he takes his first graduate course in atomic or nuclear physics. Again the revision must be thorough, from the content of the freshman courses in physics and chemistry to the choice of topics in electromagnetic theory and dynamics. Unless the student becomes familiar with quantum theory and with the parts of classical theory relevant to quantum theory while an undergraduate, he has not the time to achieve a general understanding of modern physics as a graduate student.

Perhaps this task of compression will eventually become impossible; perhaps we shall have to give up educating physicists and resign ourselves to educating nuclear physicists or solid-state physicists or

the like. This has happened in other branches of science. It would inevitably happen here if the task were simply that of compression of a growing mass of disparate observed facts. But physics has so far been fortunate enough to have the scope of its theoretical constructs expand nearly as rapidly as the volume of its experimental data. Thus it has so far not been necessary, for example, to have to teach two subjects in the time previously devoted to one; it has instead been possible to teach the concepts of a new theory which encompasses both of the earlier separate subjects, concentrating on conceptual understanding and relegating details to later specialist subjects. The range of coverage of each course has increased; by and large the number of concepts has not.

Of course the newer, more-inclusive theories embody more-sophisticated concepts and a wider range of mathematical techniques. So it is not an easy job to make them understandable to the undergraduate. However, the task is not yet impossible, in this writer's opinion.

Statistical physics is a case in point. It impinges on nearly all of modern physics and is basic for the understanding of many aspects, both experimental and theoretical. Classical thermodynamics is inadequate for many of the applications; to be useful in the areas of present active research it must be combined with quantum theory via the concepts of statistical mechanics. And this enrichment of thermodynamics should be included in the undergraduate course, so that the student can apply it in his graduate courses from the first.

Such a course needs thorough replanning, both as to choice of material and order. Topics must be omitted to make room for new items to be added. The problem is not so much the number of different concepts to be taught as the abstractness and sophistication of these concepts. The thermodynamic part has to be compressed, of course, but not at the price of excluding all variables except  $P$ ,  $V$ , and  $T$ . Engineering applications have to be omitted and special cases can be relegated to problems. Because of its importance in statistical mechanics, entropy should be stressed more than might be necessary in a course of thermodynamics alone. The kinetic theory part can be used as an introduction to the concepts of statistical mechanics, tying the material together with the Boltzmann equation, which recently became important in plasma physics. In statistical mechanics the effort must be to unify the point of view, so that each new aspect does not seem like a totally new subject. More needs to be done than is embodied in this text, but the class response so far has indicated that ensembles and partition functions are not necessarily beyond the undergraduate.

As usual, the problem is best solved experimentally, by trying out various ways of presentation to see which ones lead the student most easily to the correct point of view—which exposition brings the sparks of interest and understanding from the class. The problem would not be soluble, of course, if the students do not already possess some

knowledge of atomic structure and the basic concepts of quantum mechanics provided by a prior course in atomic structure that did not stop at Bohr orbits.

The author has been working at this pedagogical problem, off and on, for about five years, first taking recitation sections and, for the past two years, giving the lectures in the spring-semester senior course in thermodynamics and statistical mechanics for physics majors at the Institute. The present text is based on the lecture notes of the past term. Obviously the presentation has not yet reached its final form, ready for embalming in hard covers. But the results so far are encouraging, both in regard to interest aroused in the students and to concepts assimilated (at least long enough to use them in the final examination!). It is rated as a stiff course, but in the writer's experience this has never hurt a course's popularity; boredom is shunned, not work.

The writer is indebted to the subjects of his experimentation, the roughly 300 students who have attended his lectures during the past two years. Their interest, their questions, and their answers on examination papers have materially influenced the choice of subject matter and its manner of presentation. He is also particularly indebted to Professors L. Tisza and L. C. Bradley, who have read and commented on various parts of the material, and to Mr. Larry Zamick who has painstakingly checked over the present text. They are to be thanked for numerous improvements and corrections; they should not be blamed for the shortcomings that remain.

PHILIP M. MORSE

*Cambridge, Massachusetts*  
*July 1961*

# Contents

## Preface

<b>PART I. THERMODYNAMICS</b>	<b>1</b>
1. Introduction	3
Historical. Thermodynamics and statistical mechanics.	
Equilibrium states	
2. Heat, Temperature, and Pressure	7
Temperature. Heat and energy. Pressure and atomic motion. Pressure in a perfect gas	
3. State Variables and Equations of State	13
Extensive and intensive variables. Pairs of mechanical variables. The equation of state of a gas. Other equations of state. Partial derivatives	
4. The First Law of Thermodynamics	23
Work and internal energy. Heat and internal energy. Quasistatic processes. Heat capacities. Isothermal and adiabatic processes	
5. The Second Law of Thermodynamics	32
Heat engines. Carnot cycles. Statements of the second law. The thermodynamic temperature scale	
6. Entropy	40
A thermal-state variable. Reversible processes. Irreversible processes. Entropy of a perfect gas. The Joule experiment. Entropy of a gas. Entropy of mixing	
7. Simple Thermodynamic Systems	51
The Joule-Thomson experiment. Black-body radiation. Paramagnetic gas	

8. The Thermodynamic Potentials	59
The internal energy. Enthalpy. The Helmholtz and Gibbs functions. Procedures for calculation. Examples and useful formulas	
9. Changes of Phase	68
The solid state. Melting. Clausius-Clapeyron equation. Evaporation. Triple point and critical point	
10. Chemical Reactions	78
Chemical equations. Heat evolved by the reaction. Reactions in gases. Electrochemical processes	
<b>PART II. KINETIC THEORY</b>	
11. Probability and Distribution Functions	87
Probability. Binomial distribution. Random walk. The Poisson distribution. The normal distribution	
12. Velocity Distributions	96
Momentum distribution for a gas. The Maxwell distribution. Mean values. Collisions between gas molecules	
13. The Maxwell-Boltzmann Distribution	103
Phase space. The Boltzmann equation. A simple example. A more general distribution function. Mean energy per degree of freedom. A simple crystal model. Magnetization and Curie's law	
14. Transport Phenomena	116
An appropriate collision function. Electric conductivity in a gas. Drift velocity. Diffusion	
15. Fluctuations	122
Equipartition of energy. Mean-square velocity. Fluctuations of simple systems. Density fluctuations in a gas. Brownian motion. Random walk. The Langevin equation. The Fokker-Planck equations	
<b>PART III. STATISTICAL MECHANICS</b>	
16. Ensembles and Distribution Functions	137
Distribution functions in phase space. Liouville's theorem. Quantum states and phase space	
17. Entropy and Ensembles	147
Entropy and information. Information theory. Entropy for equilibrium states. Application to a perfect gas	

18. The Microcanonical Ensemble	154
Example of a simple crystal. Microcanonical ensemble for a perfect gas. The Maxwell distribution	
19. The Canonical Ensemble	163
Solving for the distribution function. General properties of the canonical ensemble. The effects of quantization. The high-temperature limit	
20. Statistical Mechanics of a Crystal	170
Normal modes of crystal vibration. Quantum states for the normal modes. Summing over the normal modes. The Debye formulas. Comparison with experiment	
21. Statistical Mechanics of a Gas	181
Factoring the partition function. The translational factor. The indistinguishability of molecules. Counting the system states. The classical correction factor. The effects of molecular interaction. The Van der Waals equation of state	
22. A Gas of Diatomic Molecules	195
The rotational factor. The gas at moderate temperatures. The vibrational factor	
23. The Grand Canonical Ensemble	202
An ensemble with variable N. The grand partition function. The perfect gas once more. Density fluctuations in a gas	
24. Quantum Statistics	208
Occupation numbers. Maxwell-Boltzmann particles. Bosons and fermions. Comparison among the three statistics. Distribution functions and fluctuations	
25. Bose-Einstein Statistics	220
General properties of a boson gas. Classical statistics of black-body radiation. Statistical mechanics of a photon gas. Statistical properties of a photon gas. Statistical mechanics of a boson gas. Thermal properties of a boson gas. The degenerate boson gas	
26. Fermi-Dirac Statistics	223
General properties of a fermion gas. The degenerate fermion gas. Behavior at intermediate temperatures	
27. Quantum Statistics for Complex Systems	238
Wave functions and statistics. Symmetric wave functions. Antisymmetric wave functions. Wave functions and equilibrium states. Electrons in a metal. Ortho- and para-hydrogen	

## CONTENTS

xi

References	249
Problems	250
Constants	268
Glossary	270
Index	272



# I

## THERMODYNAMICS



# 1 Introduction

The subject matter of this book, thermodynamics, kinetic theory, and statistical mechanics, constitutes the portion of physics having to do with heat. Thus it may be called *thermal physics*. Since heat and other related properties, such as pressure and temperature, are characteristics of aggregates of matter, the subject constitutes a part of the physics of matter in bulk. Its theoretical models, devised to correspond with the observed behavior of matter in bulk (and to predict other unobserved behavior), rely on the mathematics of probability and statistics; thus the subject may alternately be called *statistical physics*.

## Historical

The part of the subject called thermodynamics has had a long and controversial history. It began at the start of the industrial revolution, when it became important to understand the relation between heat and chemical transformations and the conversion of heat into mechanical energy. At that time the atomic nature of matter was not yet understood and the mathematical model, developed to represent the relation between thermal and mechanical behavior, had to be put together with the guidance of the crude experiments that could then be made. Many initial mistakes were made and the theory had to be drastically revised several times. This theory, now called thermodynamics, concerns itself solely with the macroscopic properties of aggregated matter, such as temperature and pressure, and their interrelations, without reference to the underlying atomic structure. The general pattern of these interrelations is summarized in the laws of thermodynamics, from which one can predict the complete thermal behavior of any substance, given a relatively few empirical relationships, obtained by macroscopic measurements made on the substance in question.

In the latter half of the nineteenth century, when the atomic nature of matter began to be understood, efforts were made to learn how the

macroscopic properties of matter, dealt with by thermodynamics, could depend on the assumed behavior of constituent atoms. The first successes of this work were concerned with gases, where the interactions between the atomic components were minimal. The results provide a means of expressing the pressure, temperature, and other macroscopic properties of the gas in terms of average values of properties of the molecules, such as their kinetic energy. This part of the subject came to be called *kinetic theory*.

In the meantime a much more ambitious effort was begun by Gibbs in this country, and by Boltzmann and others in Europe, to provide a statistical correspondence between the atomic substructure of any piece of matter and its macroscopic behavior. Gibbs called this theory *statistical mechanics*. Despite the fragmentary knowledge of atomic physics at the time, statistical mechanics was surprisingly successful from the first. Since then, of course, increased atomic knowledge has enabled us to clarify its basic principles and extend its techniques. It now provides us with a means of understanding the laws of thermodynamics and of predicting the various relations between thermodynamic variables, hitherto obtained empirically.

### Thermodynamics and Statistical Mechanics

Thus thermodynamics and statistical mechanics are mutually complementary. For example, if the functional relationship between the pressure of a gas, its temperature, and the volume it occupies is known, and if the dependence of the heat capacity of the gas on its temperature and pressure has been determined, then thermodynamics can predict how the temperature and pressure are related when the gas is isolated thermally, or how much heat it will liberate when compressed at constant temperature. Statistical mechanics, on the other hand, seeks to derive the functional relation between pressure, volume, and temperature, and also the behavior of the heat capacity, in terms of the properties of the molecules that make up the gas.

In this volume we shall first take up thermodynamics, because it is more obviously related to the gross physical properties we wish to study. But we shall continue to refer back to the underlying micro-structure, by now well understood, to remind ourselves that the thermodynamic variables are just another manifestation of atomic behavior. In fact, because it does not make use of atomic concepts, thermodynamics is a rather abstract subject, employing sophisticated concepts, which have many logical interconnections; it is not easy to understand one part until one understands the whole. In such a case it is better pedagogy to depart from strict logical presentation. Hence several derivations and definitions will be given in steps, first presented in simple form and, only after other concepts have been introduced, later re-enunciated in final, accurate form.

Part of the difficulty comes from the fact, more apparent now than earlier, that the thermodynamic quantities such as temperature and pressure are aggregate effects of related atomic properties. In thermodynamics we assume, with considerable empirical justification, that whenever a given amount of gas, in a container of given volume, is brought to a given temperature, its pressure and other thermodynamic properties will take on specific values, no matter what has been done to the gas previously. By this we do not mean that when the gas is brought back to the same temperature each molecule of the gas returns to the same position and velocity it had previously. All we mean is that the average effects of all the atoms return to their original values, that even if a particular molecule does not return to its previous position or velocity, its place will have been taken by another, so that the aggregate effect is the same.

To thus assume that the state of a given collection of atoms can be at all adequately determined by specifying the values of a small number of macroscopic variables, such as temperature and pressure, would at first seem to be an unworkable oversimplification. Even if there were a large number of different configurations of atomic positions and motions which resulted in the same measurement of temperature, for example, there is no *a priori* reason that all, or even most, of these same configurations would produce the same pressure. What must happen (and what innumerable experiments show does happen) is that a large number of these configurations do produce the same pressure and that thermodynamics has a method of distinguishing this subset of configurations from others, which do not produce the same pressure. The distinguishing feature of the favored subset is embodied in the concept of the equilibrium state.

### Equilibrium States

A detailed specification of the position, velocity, and quantum state of each atom in a given system is called a *microstate* of the system. The definition is useful conceptually, not experimentally, for we cannot determine by observation just what microstate a system is in at some instant, and we would not need to do so even if we could. As we said before, many different microstates will produce the same macroscopic effects; all we need to do is to find a method of confining our attention to that set of microstates which exhibits the simple relations between macroscopic variables, with which thermodynamics concerns itself.

Consider for a moment all those microstates of a particular gas for which the total kinetic energy of all the molecules is equal to some value  $U$ . Some of the microstates will correspond to the gas being in a state of turbulence, some parts of the gas having net momentum in one direction, some in another. But a very large number of micro-

states will correspond to a fairly uniform distribution of molecular kinetic energies and directions of motion, over all regions occupied by the gas. In these states, which we shall call the equilibrium microstates, we shall find that the temperature and pressure are fairly uniform throughout the gas. It is a fact, verified by many experiments, that if a gas is started in a microstate corresponding to turbulence, it will sooner or later reach one of the equilibrium microstates, in which temperature and pressure are uniform. From then on, although the system will change from microstate to microstate as the molecules move about and collide, it will confine itself to equilibrium microstates. To put it in other language, although the gas may start in a state of turbulence, if it is left alone long enough internal friction will bring it to that state of thermodynamic quiescence we call *equilibrium*, where it will remain.

Classical thermodynamics only deals with *equilibrium states* of a system, each of which corresponds to a set of indistinguishable microstates, indistinguishable because the temperature, the pressure, and all the other applicable thermodynamic variables have the same values for each microstate of the set. These equilibrium states are reached by letting the system settle down long enough so that quantities such as temperature and pressure become uniform throughout, so that the system has a chance to forget its past history, so to speak.

Quantities, such as pressure and temperature, which return to the same values whenever the system returns to the same equilibrium state, are called *state variables*. A thermodynamic state of a given system is thus completely defined by specifying the values of a relatively few state variables (which then become the *independent variables*), whereupon the values of all the other applicable state variables (the *dependent variables*) are uniquely determined. Dependent state variable  $U$  is thus specified as a function  $U(x, y, \dots, z)$  of the independent variables  $x, \dots, z$ , where it is tacitly understood that the functional relationship only holds for equilibrium states.

It should be emphasized that the fact that there are equilibrium states, to which matter in bulk tends to approach spontaneously if left to itself, and the fact that there are thermodynamic variables which are uniquely specified by the equilibrium state (independent of the past history of the system) are *not* conclusions deduced logically from some philosophical first principles. They are conclusions ineluctably drawn from more than two centuries of experiments.

# 2

# Heat, Temperature, and Pressure

In our introductory remarks we have used the words temperature and pressure without definition, because we could be sure the reader had encountered them earlier. Before long, however, these quantities must be defined, by describing how they are measured and also by indicating how they are related to each other and to all the other quantities that enter into the theoretical construct we call thermodynamics. As mentioned earlier, this construct is so tightly knit that an adequate definition of temperature involves other concepts and quantities, themselves defined in terms of temperature. Thus a stepwise procedure is required.

## Temperature

The first step in a definition of temperature, for example, is to refer to our natural perception of heat and cold, noting that the property seems to be one-dimensional, in that we can arrange objects in a one-dimensional sequence, from coldest to hottest. Next we note that many materials, notably gases and some liquids, perceptibly expand when heated; so we can devise and arbitrarily calibrate a thermometer. The usual calibration, corresponding to the centigrade scale, sets  $0^\circ$  at the temperature of melting ice and  $100^\circ$  at the temperature of boiling water, and makes the intermediate scale proportional to the expansion of mercury within this range. We use such a thermometer to provide a preliminary definition of temperature,  $T$ , until we have learned enough thermodynamics to understand a better one.

Next we note, by experimenting on our own or by taking the word of other experimenters, that when a body comes to equilibrium its temperature is uniform throughout. In fact we soon come to use uniformity of temperature as an experimental criterion of equilibrium.

When we turn to the question of the cause of change of temperature the usual answer, that a rise in temperature is caused by the addition of heat, has no scientific meaning until we define heat. And thus we

arrive at the source of many of the mistakes made in the early development of thermodynamics. Heat is usually measured by measuring the rise in temperature of the body to which the heat is added. This sounds like a circular definition: The cause of temperature rise is heat, which is measured by the temperature rise it causes. Actually it is something more, for it assumes that there is something unique called heat, which produces the same temperature change no matter how the heat is produced. Heat can be generated by combustion, by doing work against friction, by sending electric current through a resistor, or by rapidly compressing a gas. The amount of heat produced in each of these ways can be measured in terms of the temperature rise of a standard body and the effects are proportional; twice the combustion producing twice the temperature rise, for example.

The early measurements of heat were all consistent with the theory that heat is a "fluid" similar to the "electric fluid" which was also being investigated at that time. The heat fluid was supposed to be bound to some atoms in the body; it could be detached by pressure, friction, or combustion; in its free state it would affect thermometers. It seemed at first that the amount  $Q$  of "free heat fluid" present in a body should be a thermodynamic state variable such as pressure or temperature, a definite function of the independent variables that define the equilibrium state. The  $Q$  for a particular amount of gas was supposed to be a specific function of the temperature and pressure of the gas, for instance.

Later, however, it was demonstrated that heat is just one manifestation of the energy possessed by a body, that heat could be transformed into mechanical energy and vice versa. Historically, this change in theory is reflected by the change in the units used to measure heat. At first the unit was the kilogram-calorie, the amount of heat required to raise a kilogram of water from 4 to 5°C. More recently heat has been measured in terms of the usual units of energy, the *joule* in the mks system of units. Careful measurement of the energy lost to friction, or that lost in passing current through a resistor, together with the resulting temperature rise in water placed in thermal contact, shows that a kilogram-calorie of heat is equal to 4182 joules.

### Heat and Energy

As soon as we realize that heat is just a particular manifestation of the energy content of a body, we see that  $Q$  *cannot be a state variable*. For we can add energy to a body in the form of heat and then take it away in the form of mechanical energy, bringing the body back to its initial equilibrium state at the end of such a cycle of operation. If heat were a state variable, as much would have to be given off during the cycle as was absorbed for  $Q$  to come back to its original value at the

end of the cycle. But if heat can be changed to work, the net amount of heat added during the cycle may be positive, zero, or negative, depending on the net amount of work done by or on the body during the cycle. The quantity which is conserved, and which thus is the state variable, is  $U$ , the *energy* possessed by the body, which can be drawn off either as heat or as mechanical work, depending on the circumstances. As we shall see more clearly later, heat represents that energy content of the body which is added or removed in disorganized form; work is the energy added or removed in organized form; within certain limits disorganization can be changed to organization and vice versa.

We can increase the energy of a body by elevating it, doing work against the force of gravity. This increase is in potential energy, which is immediately available again as work. The temperature of the body is not changed by the elevation, so its heat content is not changed. We can translate the potential energy into organized kinetic energy by dropping the body; this also makes no change in its heat content. But if the body, in its fall, hits the ground, the organized motion of the body is changed to disorganized internal vibrations; its temperature rises; heat has been produced. In a sense, the reason that classical thermodynamics is usually limited to a study of equilibrium states is because an equilibrium state is the state in which heat energy can easily and unmistakably be distinguished from mechanical energy. Before equilibrium is reached, sound waves or turbulence may be present, and it is difficult to decide when such motion ceases to be "mechanical" and becomes "thermal."

A preliminary definition of pressure can be more quickly achieved, for pressure is a mechanical quantity; we have encountered it in hydromechanics and acoustics. Pressure is a simple representative of a great number of internal stresses which can be imposed on a body, such as tensions or torques or shears, changes in any of which represent an addition or subtraction of mechanical energy to the body. Pressure is more usually encountered in thermodynamic problems; it is the only stress that a gas can sustain in equilibrium. The usual units of pressure are *newtons per square meter*, although the *atmosphere* (about  $10^5$  newtons per square meter) is also used.

### Pressure and Atomic Motion

The pressure exerted by a gas on its container walls is a very good example of a mechanical quantity which is the resultant of the random motions of the gas molecules and which nonetheless is a remarkably stable function of a relatively small number of state variables. To illustrate this point we shall digress for a few pages into a discussion of kinetic theory. The pressure  $P$  on a container wall is the force exerted by the gas, normal to an area  $dA$  of the wall, divided by  $dA$ . This force is caused by the collisions of the gas molecules against the

area  $dA$ . Each collision delivers a small amount of momentum to  $dA$ ; the amount of momentum delivered per second is the force  $P dA$ .

Let us assume a very simplified model of a gas, one consisting of  $N$  similar atoms, each of mass  $m$  and of "negligible" dimensions, with negligible interactions between them so that the sole energy of the  $i$ th atom is its kinetic energy of translation,  $(1/2)m(v_{ix}^2 + v_{iy}^2 + v_{iz}^2)$ . The gas is confined in a container of internal volume  $V$ , the walls of which are perfect reflectors for incident gas atoms. By "negligible dimensions" we mean the atoms are very small compared to the mean distance of separation, so collisions are very rare and most of the time each atom is in free motion. We also mean that we do not need to consider the effects of atomic rotation. We shall call this simple model a *perfect gas of point atoms*.

Next we must ask what distribution of velocities and positions of the  $N$  atoms in volume  $V$  corresponds to a state of equilibrium. As the atoms rebound from the walls and from each other, they cannot lose energy, for the collisions are elastic. In collisions between the atoms what energy one loses the other gains. The total energy of translational motion of all the atoms,

$$U = \frac{1}{2}m \sum_{N=1}^N (v_{ix}^2 + v_{iy}^2 + v_{iz}^2) = \frac{1}{2}m \sum_i v_i^2 = N \langle \text{K.E.} \rangle_{\text{tran}} \quad (2-1)$$

is constant. The last part of this set of equations defines the average kinetic energy  $\langle \text{K.E.} \rangle_{\text{tran}}$  of translation of an atom of the gas as being  $U$  divided by the number of atoms  $N$ . (The angular brackets  $\langle \rangle$  will symbolize average values.)

As the gas settles down to equilibrium,  $U$  does not change but the randomness of the atomic motion increases. At equilibrium the atoms will be uniformly distributed throughout the container, with a density  $(N/V)$  atoms per unit volume; their velocities will also be randomly distributed, as many moving in one direction as in another. Some atoms are going slowly and some rapidly, of course, but at equilibrium the total  $x$  component of atomic momentum,  $\sum m v_{ix}$ , is zero; similarly with the total  $y$  and  $z$  components of momentum. The total  $x$  component of the kinetic energy,  $(1/2)\sum m v_{ix}^2$ , is not zero, however. At equilibrium it is equal to the total  $y$  component and to the total  $z$  component, each of which is equal to one-third of the total kinetic energy, according to Eq. (2-1),

$$\sum_i \frac{1}{2} m v_{ix}^2 = \frac{1}{3} N \langle \text{K.E.} \rangle_{\text{tran}} \quad (2-2)$$

At equilibrium all directions of motion are equally likely.

### Pressure in a Perfect Gas

Next we ask how many atoms strike the area  $dA$  of container wall per second. For simplicity we orient the axes so that the positive  $x$  axis points normally into  $dA$  (Fig. 2-1). Consider first all those atoms

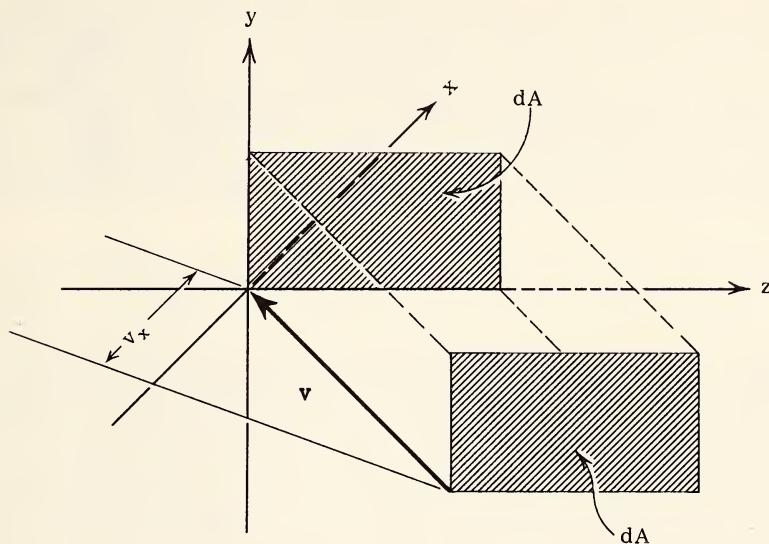


FIG. 2-1. Reflection of atoms, with velocity  $v$ , from area element  $dA$  in the  $yz$  plane.

in  $V$  which have their  $x$  component of velocity,  $v_{ix}$ , equal to some value  $v_x$  ( $v_x$  must be positive if the atom is to hit  $dA$ ). All these kinds of atoms, which are a distance  $v_x$  from the wall or closer, will hit the wall in the next second, and a fraction proportional to  $dA$  of those will hit  $dA$  in the next second. In fact a fraction  $(v_x/V) dA$  of all the atoms in  $V$  which have  $x$  component of velocity equal to  $v_x$  will hit  $dA$  per second. Each of these atoms, as it strikes  $dA$  during the second, rebounds with an  $x$  component  $-v_x$ , so each of these atoms imparts a momentum  $2mv_x$  to  $dA$ . Thus the momentum imparted per second by the atoms with  $x$  velocity equal to  $v_x$  is  $2mv_x(v_x/V) dA$  times the total number of atoms in  $V$  having  $x$  velocity equal to  $v_x$ . And therefore the total momentum imparted to  $dA$  per second is the sum of  $(2mv_{ix}^2/V) dA$  for each atom in  $V$  that has a positive value of  $v_{ix}$ .

Since half the atoms have a positive value of  $v_{ix}$ , the total momentum imparted to  $dA$  per second is

$$\frac{1}{2} \sum_{i=1}^N (2mv_{ix}^2/V) dA = \frac{2}{3} (N/V) \langle K.E. \rangle_{tran} dA = \frac{2}{3} (U/V) dA \quad (2-3)$$

where we have used Eq. (2-2) to express the result in terms of the mean atomic kinetic energy, defined in Eq. (2-1). Since this is equal to the force  $P dA$  on  $dA$ , we finally arrive at an equation relating the pressure  $P$  of a perfect gas of point atoms in a volume  $V$ , in terms of the mean kinetic energy  $\langle K.E. \rangle_{tran}$  of translation per atom or in terms of the total energy content of the gas per unit volume  $(U/V)$  (total as long as the only energy is kinetic energy of translation, that is):

$$P = \frac{2}{3} (U/V) \quad \text{or} \quad PV = \frac{2}{3} U = \frac{2}{3} N \langle K.E. \rangle_{tran} \quad (2-4)$$

This is a very interesting result, for it demonstrates the great stability of the relationships between aggregate quantities such as  $P$  and  $U$  for systems in equilibrium. The relationship of Eq. (2-4) holds no matter what the distribution in speed the atoms have as long as their total energy is  $U$ , as long as the atoms are uniformly distributed in space, and as long as all directions of motion are equally likely (i.e., as long as the gas is in equilibrium). Subject to these provisos, every atom could have kinetic energy  $\langle K.E. \rangle_{tran}$ , or half of them could have kinetic energy  $(1/2) \langle K.E. \rangle_{tran}$  and the other half energy  $(3/2) \langle K.E. \rangle_{tran}$ , or any other distribution having an average value  $\langle K.E. \rangle_{tran}$ . As long as it is uniform in space and isotropic in direction the relation between  $P$ ,  $V$ , and  $N \langle K.E. \rangle_{tran}$  is that given in Eq. (2-4). Even the proportionality constant is fixed;  $PV$  is not just proportional to  $N \langle K.E. \rangle_{tran}$  — the factor is  $2/3$ , no matter what the velocity distribution is.

From our earlier discussion we may suspect that  $\langle K.E. \rangle_{tran}$  is a function of the gas temperature  $T$ ; if  $T$  is increased, the kinetic energy of the gas atoms should increase. We shall see what this relationship is in the next chapter.

# 3

## State Variables and Equations of State

To recapitulate, when a thermodynamic system is in a certain equilibrium state, certain aggregate properties of the system, such as pressure and temperature, called *state variables*, have specific values, determined only by the state and not by the previous history of the system. Alternately, specifying the values of a certain number of state variables specifies the state of the system; the number of variables, required to specify the state uniquely, depends on the system and on its constraints. For example, if the system is a definite number of like molecules in gaseous form within a container, then only two variables are needed to specify the state—either the pressure of the gas and the volume of the container, or the pressure and temperature of the gas, or else the volume and temperature. If the system is a mixture of two gases (such as hydrogen and oxygen) which react chemically to form a third product (such as water vapor), the relative abundance of two of the three possible molecular types must be specified, in addition to the total pressure and volume (or  $P$  and  $T$ , or  $T$  and  $V$ ), to determine the state. If the gas is paramagnetic, and we wish to investigate its thermomagnetic properties, then the strength  $\mathcal{C}$  of the applied magnetic field (or else the magnetic polarization of the gas) must also be specified.

### Extensive and Intensive Variables

One state variable is simply the amount of each chemical substance present in the system. The convenient unit for this variable is the *mole*; 1 mole of a substance, which has molecular weight  $M$ , is  $M$  kilograms of the substance (1 mole of hydrogen gas is 2 kg of  $H_2$ , 1 mole of oxygen is 32 kg of  $O_2$ ). By definition, each mole contains the same number of molecules,  $N_0 = 6 \times 10^{26}$ , called *Avogadro's number*. In many respects a mole of gas behaves the same, no matter what its composition. When the thermodynamic system is made up of a single substance then the number of moles present (which we shall denote by

the letter  $n$ ) is constant. But if the system is a chemically reacting mixture the  $n$ 's for the different substances may change.

State variables are of two sorts, one sort directly proportional to  $n$ , the other not. For example, suppose we have two equal amounts of the same kind of gas, each of  $n$  moles and each in equilibrium at the same temperature  $T$  in containers of equal volume  $V$ . We then connect the containers so the two samples of gas can mix. The combined system now has  $2n$  moles of gas in a volume  $2V$ , and the total internal energy of the system is twice the internal energy  $U$  of each original part. But the common temperature  $T$  and pressure  $P$  of the mixed gas have the same values they had in the original separated states. Variables of the former type, proportional to  $n$  (such as  $U$  and  $V$ ), are called *extensive* variables; those of the latter type (such as  $T$  and  $P$ ) are called *intensive* variables. At thermodynamic equilibrium the intensive variables have uniform values throughout the system.

A basic state variable for all thermodynamic systems (almost by definition) is its temperature  $T$ , which is an intensive variable. At present we have agreed to measure its value by a thermometer; a better definition will be given later. Related to the temperature is the heat capacity of the system, the heat required to raise the system 1 degree in temperature. Because heat is not a state variable, the amount of heat required depends on the way the heat is added. For example, the amount of heat required to raise  $T$  by 1 degree, when the volume occupied by the system is kept constant, is called the *heat*

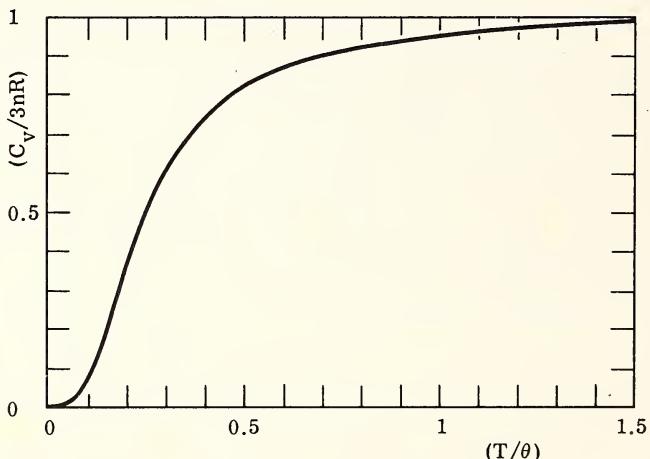


FIG. 3-1. Heat capacity at constant volume  $C_V$  of a solid, in units of  $3nR$ , where  $R$  is the gas constant [see Eq. (3-1)] as a function of temperature, in units of the characteristic temperature  $\theta$  of the solid (see Fig. 20-1).

*capacity at constant volume* and is denoted by  $C_V$ . The heat required to raise  $T$  1 degree when the pressure is kept constant is called the *heat capacity at constant pressure* and is denoted  $C_P$ . A system at constant pressure usually expands when heated, thus doing work, so  $C_P$  is usually greater than  $C_V$ .

These heat capacities are state variables, in fact they are extensive variables; their units are joules per degree. The capacities per mole of substance,  $c_V = (C_V/n)$  and  $c_P = (C_P/n)$ , are called *specific heats*, at constant volume or pressure, respectively. They have been measured, for many materials, and a number of interesting regularities have emerged. For example, the specific heat at constant volume,  $c_V$ , for any monatomic gas is roughly equal to 12,000 joules per degree centigrade per mole, independent of  $T$  and  $P$  over a wide range of states, whereas  $c_V$  for diatomic gases is roughly 20,000 joules per degree centigrade per mole, with a few exceptions. A typical plot of  $c_V$  for a solid is given in Fig. 3-1, showing that  $c_V$  is independent of  $T$  for solids only when  $T$  is quite large (there are more exceptions to this rule with solids than with gases).

### Pairs of Mechanical Variables

Other state variables are of a mechanical, rather than thermal, type. For example, there is the pressure  $P$  (in newtons per square meter), an intensive variable appropriate for fluids, although applicable also for solids that are uniformly compressed (in general, in solids, one needs a tensor to describe the stress). Related to  $P$  is the extensive variable  $V$  (in cubic meters), the volume occupied by the system. The pair define a mechanical energy; work  $P dV$  (in joules) is done by the system on the container walls if its volume is increased by  $dV$  when it is in equilibrium at pressure  $P$ . The pair  $P$  and  $V$  are the most familiar of the mechanical state variables. For a bar of material, the change in dimensions may be simple stretching, in which case the extensive variable would be the length  $L$  of the bar, the intensive variable would be the tension  $J$ , and the work done on the bar by stretching it an additional amount  $dL$  would be  $J dL$ .

Or possibly the material may be polarized by a magnetic field. The intensive variable here is the impressed magnetic intensity  $\mathcal{H}$  (in ampere turns per meter) and the extensive variable is the total magnetization  $\mathfrak{M}$  of the material. Reference to a text on electromagnetic theory will remind us that, related to  $\mathcal{H}$ , there is the magnetic induction field  $\mathfrak{B}$  (in webers per square meter). For paramagnetic material of magnetic susceptibility  $\chi$ , a magnetic field causes a polarization  $\mathfrak{P}$  of the material, which is related to  $\mathcal{H}$  and  $\mathfrak{B}$  through the equation  $\mathfrak{B} = \mu_0 \mathcal{H}(1 + \chi) = \mu_0(\mathcal{H} + \mathfrak{P})$ , where  $\mu_0$  is the permeability of vacuum  $4\pi \times 10^{-7}$  henrys per meter. The total energy contributed to the material occupying volume  $V$ , by application of field  $\mathcal{H}$ , is  $(1/2) \mu_0 V \mathcal{H} \mathfrak{P}$ ,

exclusive of the "energy of the vacuum"  $(1/2)\mu_0V\mathcal{H}^2$ , the total magnetic energy being  $(1/2)\mathcal{H}\mathcal{G}V$ .

Suppose we define the total *magnetization* of the body as being the quantity  $\mathfrak{M} = \mu_0V\varphi$  (in weber-meters); for paramagnetic materials  $\mathfrak{M}$  would equal  $\mu_0V\chi\mathcal{H}$ . Then the magnetic work done on the body in magnetic field  $\mathcal{H}$ , when its magnetization is increased by  $d\mathfrak{M}$ , would be  $\mathcal{H}d\mathfrak{M}$ , the integral of which, for  $\mathfrak{M} = \mu_0V\chi\mathcal{H}$ , becomes  $(1/2)\mu_0V\mathcal{H}\varphi$ , as desired. Magnetization  $\mathfrak{M}$  is thus the extensive variable related to  $\mathcal{H}$ . A similar pair of state variables can be devised for dielectrics and the electric field.

There is also an intensive variable, related to the variable  $n$ , the number of moles of material in the system. If we add  $dn$  moles of material to the system we add energy  $\mu dn$ , where  $\mu$  is called the *chemical potential* of the material. Its value can be measured by measuring the heat generated by a chemical reaction, as will be shown in Chapter 10.

As we mentioned earlier, we need to determine experimentally a certain minimum number of relationships between the state variables of a system before the theoretical machinery of thermodynamics can "take hold" to predict the system's other thermal properties. One of these relationships is the dependence of one of the heat capacities, either  $C_V$  or  $C_p$  (or  $C_L$  or  $C_J$ , or  $C_{\mathcal{H}}$  or  $C_{\mathfrak{M}}$  depending on the mechanical or electromagnetic variables of interest) on  $T$  and on  $P$  or  $V$  (or  $L$  or  $J$ , or on  $\mathcal{H}$  or  $\mathfrak{M}$ , as the case may be). We shall show later that, if  $C_V$  is measured,  $C_p$  can be computed (and similarly for the other pairs of heat capacities); thus only one heat capacity needs to be measured as a function of the independent variables of the system.

### The Equation of State of a Gas

Another necessary empirical relationship is the relation between the pair of mechanical variables  $P$  and  $V$  (or  $J$  and  $L$ , or  $\mathcal{H}$  and  $\mathfrak{M}$ ) and the temperature  $T$  for the system under study. Such a relationship, expressed as a mathematical equation, is called an *equation of state*. There must be an equation of state known for each pair of mechanical variables of interest. We shall write down some of those of general interest, which will be used often later.

Parenthetically it should be noted that although these relationships are usually experimentally determined, in principle it should be possible to compute them from our present knowledge of atomic behavior, using statistical mechanics.

The equation of state first to be experimentally determined is the one relating  $P$ ,  $V$ , and  $T$  for a gas, discovered by Boyle and by Charles. The relation is expressed by the equation  $PV = nR(T + T_0)$ , where, if  $T$  is in degrees centigrade, constant  $R$  is roughly 8300 joules per degree centigrade per mole and  $T_0$  is roughly 273 C. This

suggests that we change the origin of our temperature scale from 0°C to  $-273^{\circ}\text{C}$ . The temperature measured from this new origin (called *absolute zero*) is called absolute temperature and is expressed in *degrees Kelvin*. For T in degrees Kelvin this equation of state is

$$PV = nRT \quad (3-1)$$

Actually this is only a rough approximation of the equation of state of an actual gas. Experimentally it is found to be a pretty good approximation for monatomic gases, like helium. Moreover, the ratio  $PV/nT$  for *any* gas approaches the value 8315 joules per degree Kelvin per mole as  $P/T$  approaches zero. Since a gas is "most gassy" when the pressure is small and the atoms are far apart, we call Eq. (3-1) the *perfect gas law* and call any gas that obeys it a perfect gas. We could, alternately, use Eq. (3-1) as another (but not the final) definition of temperature; T is equal to  $PV/nR$  for a perfect gas.

We now are in a position to illustrate how kinetic theory can supplement an empirical thermodynamic formula with a physical model. In the previous chapter we calculated the pressure exerted by a gas of N point particles confined at equilibrium in a container of volume V. This should be a good model of a perfect gas. As a matter of fact Eq. (2-4) has a form remarkably like that of Eq. (3-1). All we need to do to make the equations identical is to set

$$nRT = \frac{2}{3}U = \frac{2}{3}N\langle\text{K.E.}\rangle_{\text{tran}}$$

The juxtaposition is most suggestive. We have already pointed out that N, the number of molecules, is equal to  $nN_0$ , where  $N_0$ , Avogadro's number, is equal to  $6 \times 10^{23}$  for any substance. Thus we reach the remarkably simple result, that  $RT = (2/3)N_0\langle\text{K.E.}\rangle_{\text{tran}}$  for any perfect gas. For this model, therefore, the average kinetic energy per molecule is proportional to the temperature, and the proportionality constant  $(3/2)(R/N_0)$  is independent of the molecular mass. The ratio  $(R/N_0) = 1.4 \times 10^{-23}$  joules per degree Kelvin is called k, the *Boltzmann constant*.

Thus the model suggests that for those gases which obey the perfect gas law fairly accurately, the average kinetic energy of molecular translation is directly proportional to the temperature,

$$\langle\text{K.E.}\rangle_{\text{tran}} = \frac{3}{2}kT \quad \text{perfect gas} \quad (3-2)$$

independent of the molecular mass. Only the kinetic energy of translation enters into this formula; our model of point atoms assumed their rotational kinetic energy was negligible. We might expect that this

would be true for actual monatomic gases, like helium and argon, and that for these gases the total internal energy is

$$U = N \langle K.E. \rangle_{\text{tran}} = \frac{3}{2} N k T = \frac{3}{2} n R T \quad (3-3)$$

Measurement shows this to be nearly correct [see discussion of Eq. (6-11)]. For polyatomic gases  $U$  is greater, corresponding to the additional kinetic energy of rotation (the additional term does not enter into the equation for  $P$ , however). We shall return to this point, to enlarge on and to modify it, as we learn more.

### Other Equations of State

Of course the equation of state for an actual gas is not as simple as Eq. (3-1). We could, instead of transforming our measurements into an equation, simply present the relationship between  $P$ ,  $V$ , and  $T$  in the form of a table of numbers or a set of curves. But, as we shall see, the thermal behavior of bodies usually is expressed in terms of the first and second derivatives of the equation of state, and taking derivatives of a table of numbers is tedious and subject to error. It is often better to fit an analytic formula to the data, so we can differentiate it more easily.

A formula that fits the empirical behavior of many gases, over a wider range of  $T$  and  $P$  than does Eq. (3-1), is the *Van der Waals* approximation,

$$(V - nb) \left( P + \frac{an^2}{V^2} \right) = nRT \quad (3-4)$$

For large-enough values of  $V$  this approaches the perfect gas law. Typical curves for  $P$  against  $V$  for different values of  $T$  are shown in Fig. 3-2. For temperatures smaller than  $(8a/27bR)$  there is a range of  $P$  and  $V$  for which a given value of pressure corresponds to three different values of  $V$ . This represents (as we shall see later) the transition from gas to liquid. Thus the Van der Waals formula covers approximately both the gaseous and liquid phases, although the accuracy of the formula for the liquid phase is not very good.

It is also possible to express the equation of state as a power series in  $(1/V)$ ,

$$P = (nRT/V) \left[ 1 + \frac{n}{V} B(T) + \left( \frac{n}{V} \right)^2 C(T) + \dots \right] \quad (3-5)$$

This form is called the *virial equation* and the functions  $B(T)$ ,  $C(T)$ , etc., are called virial coefficients. Values of these coefficients and

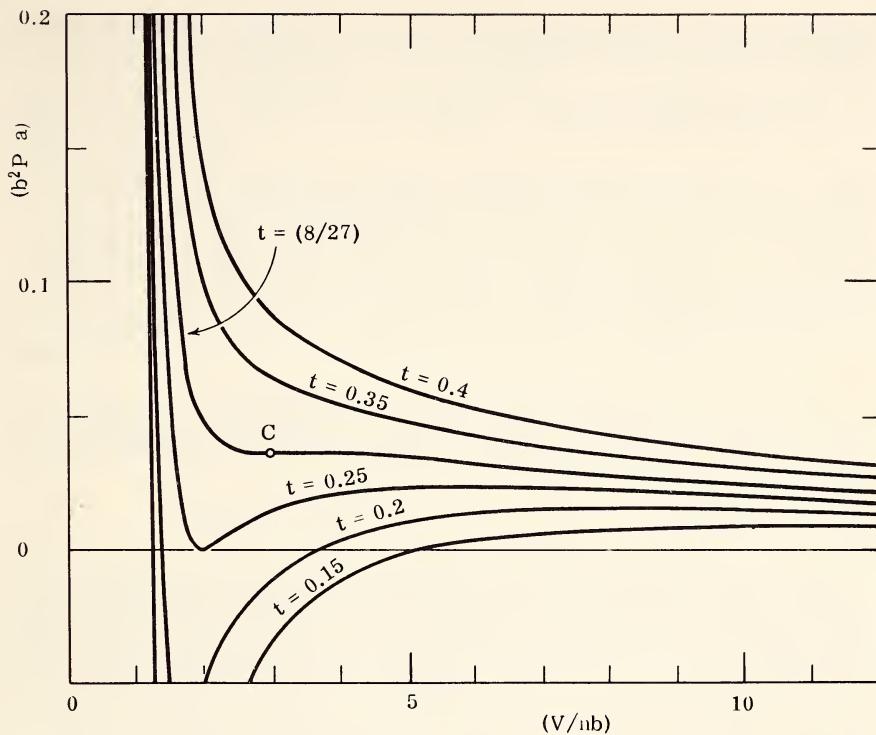


FIG. 3-2. The Van der Waals equation of state. Plots of  $b^2 P/a$  versus  $V/nb$  for different values of  $t = RbT/a$ . Point C is the critical point (see Fig. 9-3).

their derivatives can then be tabulated or plotted for the substance under study.

Corresponding equations of state can be devised for solids. A simple one, which is satisfactory for temperatures and pressures that are not too large (below the melting point for  $T$ , up to several hundred atmospheres for  $P$ ), is

$$V = V_0(1 + \beta T - \kappa P) \quad (3-6)$$

Both  $\beta$ , which is called the *thermal expansion coefficient*, and  $\kappa$ , called the *compressibility*, are small quantities, of the order of  $10^{-6}$  for metals, for example. They are not quite constant; they vary as  $T$  and  $P$  are changed, although the variation for most solids is not large for the usual range of  $T$  and  $P$ .

The other pairs of mechanical variables also have their equations of state. For example, in a stretched rod the relation between tension

$J$  and length  $L$  and temperature  $T$  is, for stretches within the elastic limit,

$$J = (A + BT)(L - L_0) \quad (3-7)$$

where  $A$ ,  $B$ , and  $L_0$  are constants (approximately);  $B$  is negative for many substances but positive for a few, such as rubber.

Likewise there is a magnetic equation of state, relating magnetic intensity  $\mathfrak{H}$ , magnetization  $\mathfrak{M}$ , and  $T$ . For paramagnetic materials, for example, *Curie's law*,

$$\mathfrak{M} = (nD\mathfrak{H}/T) \quad (3-8)$$

is a fairly good approximation, within certain limits for  $T$  and  $\mathfrak{H}$ . The Curie constant  $D$  is proportional to the magnetic susceptibility of the substance.

### Partial Derivatives

In all these equations there is a relationship between at least three variables. We can choose any pair of them to be the independent variables; the other one is then a dependent variable. We shall often wish to compute the rate of change of a dependent variable with respect to one of the independent variables, holding the other constant. This rate, called a partial derivative, is discussed at length in courses in advanced calculus. In thermodynamics, since we are all the time changing from one pair of independent variables to another, we find it advisable to label each partial by *both* independent variables, the one varied and the one held constant. The partial  $(\partial P/\partial V)_T$ , for example, is the rate of change of  $P$  with respect to  $V$ , when  $T$  is held constant;  $V$  and  $T$  are the independent variables in this case, and  $P$  is expressed explicitly as a function of  $V$  and  $T$  before performing the differentiation.

There are a number of relationships between partial derivatives that we shall find useful. If  $z$  and  $u$  are dependent variables, functions of  $x$  and  $y$ , then, by manipulation of the basic equation

$$dz = (\partial z/\partial x)_y dx + (\partial z/\partial y)_x dy$$

we can obtain

$$\left(\frac{\partial z}{\partial x}\right)_y = \frac{(\partial u/\partial x)_y}{(\partial u/\partial z)_y} = \frac{1}{(\partial x/\partial z)_y}; \quad \left(\frac{\partial x}{\partial y}\right)_z = -\frac{(\partial z/\partial y)_x}{(\partial z/\partial x)_y} \quad (3-9)$$

The last equation can be interpreted as follows: On the left we express  $x$  as a function of  $y$  and  $z$ , on the right  $z$  is expressed as a function

of  $x$  and  $y$  before differentiating and the ratio is then reconverted to be a function of  $y$  and  $z$  to effect the equation. Each partial is itself a function of the independent variables and thus may also be differentiated. Since the order of differentiation is immaterial, we have the useful relationship

$$\left[ \frac{\partial}{\partial x} \left( \frac{\partial z}{\partial y} \right)_x \right]_y = \left[ \frac{\partial}{\partial y} \left( \frac{\partial z}{\partial x} \right)_y \right]_x \quad (3-10)$$

As an example of the use of these formulas, we can find the partial  $(\partial V/\partial T)_P$ , as function of  $P$  and  $V$  or of  $T$  and  $V$ , for the Van der Waals formula (3-4):

$$\begin{aligned} \left( \frac{\partial V}{\partial T} \right)_P &= - \frac{(\partial P/\partial T)_V}{(\partial P/\partial V)_T} = \frac{nR/(V - nb)}{[nRT/(V - nb)^2] - (2an^2/V^3)} \\ &= \frac{R(V - nb)V^3}{RTV^3 - 2an(V - nb)^2} \end{aligned}$$

We shall often be given the relevant partial derivatives of a state function and be required to compute the function itself by integration. If  $(\partial z/\partial x)_y = f(x)$  and  $(\partial z/\partial y)_x = g(y)$  this is straightforward; we integrate each partial separately and add

$$z = \int f(x) dx + \int g(y) dy$$

But if either partial depends on the other independent variable it is not quite so simple. For example, if  $(\partial z/\partial x)_y = f(x) + ay$  and  $(\partial z/\partial y)_x = ax$ , then the integral is

$$z = \int f(x) dx + axy \quad [ \text{not } \int f(x) dx + 2axy ]$$

as may be seen by taking partials of  $z$ . The cross term appears in both partials and we include it only once in the integral. To thus coalesce two terms of the integral, the two terms must of course be equal. This seems to be assuming more of a relationship between  $(\partial z/\partial x)_y$  and  $(\partial z/\partial y)_x$  than we have any right to do, until we remember that they are related, according to Eq. (3-10), in just the right way so the cross terms can be coalesced. When this is so, the differential  $dz = (\partial z/\partial x)_y dx + (\partial z/\partial y)_x dy$  is a *perfect differential*, which can be integrated in the manner just illustrated to obtain  $z$ , a function of  $x$  and  $y$ , the integrated value coming out the same no matter what path in the  $x, y$  plane we choose to perform the integration along, as long as the terminal points of the path are unchanged (Fig. 3-3).

A differential  $dz = f(x, y) dx + g(x, y) dy$ , where  $(\partial f/\partial y)_x$  is not equal to  $(\partial g/\partial x)_y$ , results in an integral which depends on the path of integra-

tion as well as the end points, is called an *imperfect differential*, and is distinguished by the bar through the  $d$ . The integral of the perfect differential  $dz = y \, dx + x \, dy$ , from  $(0,0)$  to  $(a,b)$  over the path from  $(0,0)$

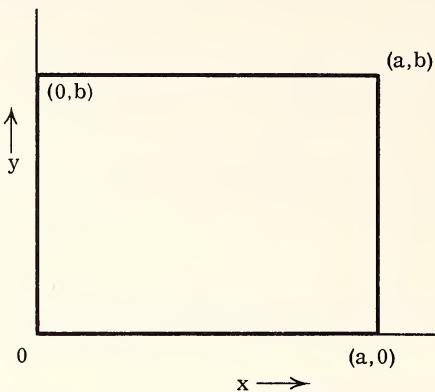


FIG. 3-3. Integration in the  $xy$  plane.

to  $(0,b)$  to  $(a,b)$  is  $0 \cdot \int_0^b dy + b \cdot \int_0^a dx = ab$ , which equals that for the path

$(0,0), (a,0), (a,b), 0 \cdot \int_0^a dx + a \cdot \int_0^b dy = ab$ . On the other hand, the integral

from  $(0,0)$  to  $(a,b)$  of the imperfect integral  $\bar{dz} = y \, dx - x \, dy$  is  $ab$  over the first route and  $-ab$  over the second. Such a differential cannot be integrated to produce a state function  $z(x,y)$ . However, we can multiply the imperfect differential  $\bar{dz}$  by an appropriate function of  $x$  and  $y$  (in this case  $1/y^2$ ), which will turn it into a perfect differential,  $du$ ; in this case

$$(\bar{dz}/y^2) = du = (1/y) \, dx - (x/y^2) \, dy \quad \text{and} \quad u = (x/y)$$

The factor that converts an imperfect differential into a perfect one is called an *integrating factor*. One always exists (although it may be hard to find) for differentials of two independent variables. For more than two independent variables there are imperfect differentials for which no integrating factor exists.

# 4

## The First Law of Thermodynamics

An important state function is the *internal energy*  $U$  of the system. For a perfect gas of point atoms, Eq. (3-3) indicates that  $U = (3/2)nRT$ , if  $T$  and  $V$ , or  $T$  and  $P$ , are the independent variables;  $U = (3/2)PV$  if  $P$  and  $V$  are. The internal energy of a system is an extensive variable.

### Work and Internal Energy

The internal energy  $U$  can be changed by having the system do work  $dW$  against some externally applied force, or by having this force do work  $-dW$  on the system. For example, if the system is confined under uniform pressure, an increase in volume would mean that the system did work  $dW = P dV$ ; if the system is under tension  $J$ , it would require work  $-dW = J dL$  to be done on the system to increase its length  $dL$ . Similarly an increase in magnetization  $dM$  in the presence of a field  $\mathcal{H}$  will increase  $U$  by  $\mathcal{H} dM$ . Or, if  $dn$  moles of a substance with chemical potential  $\mu$  is added,  $U$  would increase by  $\mu dn$ . In all these cases work  $dW$  is being done in an organized way by the system and  $U$  is increased by  $-dW$ . Note our convention, a positive  $dW$  is work done *by* the system, a negative value represents work done *on* the system, so that the change in  $U$  is opposite in sign to  $dW$ .

Note also that we have been using the symbol of the imperfect differential for  $dW$ , implying that the amount of work done by the system depends on the path (i.e., on how it is done). For example, the work done by a perfect gas in going from state 1 of Fig. 4-1 to state 2 differs whether we go via path a or path b. Along path 1a,  $V$  does not change, so no work is done by or on the gas, although the temperature changes from  $T_1 = (P_1 V_1 / nR)$  to  $T_a = (P_2 V_1 / nR)$ . Along path a<sub>2</sub>,  $P$  does not change, so that the work done by the gas in going along the whole of path 1a2 is  $\Delta W_a = P_2 (V_2 - V_1)$ . Similarly, the work done by the gas in going along path 1b2 is  $\Delta W_b = P_1 (V_2 - V_1)$ , differing from  $\Delta W_a$  by the factor  $P_1$ . This same sort of argument can be used to

show that work done by the system, in consequence of a variation of any of the mechanical variables that describe its state, cannot be a state variable.

Something more should be said about the meaning of the diagram of Fig. 4-1. In our calculations we tacitly assumed that at each point along each path the system was in an equilibrium state, for which the equation of state  $PV = nRT$  held. But for a system to be in equilibrium, so that  $P$  and  $T$  have any meaning, it must be allowed to settle down, so that  $P$  and  $V$  (and therefore  $T$ ) are assumed constant. How, then, can we talk about going along a path, about changing  $P$  and  $V$ , and at the same time assume that the system successively occupies equilibrium states as the change is made? Certainly if the change is made rapidly, sound waves and turbulence will be created and the equation of state will no longer hold. What we must do is to make the change slowly and in small steps, going from 1 to  $i$  and waiting till the system settles down, then going slowly to  $j$ , and so on. Only in the limit of many steps and slow change can we be sure that the system is never far from equilibrium and that the actual work done will approach the value computed from the equation of state. In thermodynamics we have to limit our calculations to such slow, stepwise changes (called *quasistatic* processes) in order to have our formulas hold during the change. This may seem to be an intolerable limitation on the kinds of processes thermodynamics can deal with; we shall see later that the limitation is not as severe as it may seem.

### Heat and Internal Energy

If the only way to change the system's energy is to perform work on it or have it do work, then the picture would be simple. Not only  $dU$  but also  $dW$  would be a perfect differential; whatever work was performed on the system could eventually be recovered as mechanical (or electrical or magnetic) energy. This was the original theory of thermodynamic systems; work was work and heat was heat. The introduction of heat  $dQ$  served to raise the temperature of the body (indeed the rise in temperature was the usual way in which the heat added could be measured, as was pointed out at the beginning of Chapter 2), and when the body was brought back to its initial temperature it would have given up the same heat that had been given it earlier. We could thus talk about an internal energy of the system, which was the net balance of work done on or by the system, and we could talk about the heat possessed by the body, the net balance of heat intake and output, measured by the body's temperature.

It was quite a shock to find that this model of matter in bulk was inconsistent with observation. A body's temperature could be changed by doing work on it; a body could take in heat (from a furnace, say) and produce mechanical work. It was realized that we cannot talk

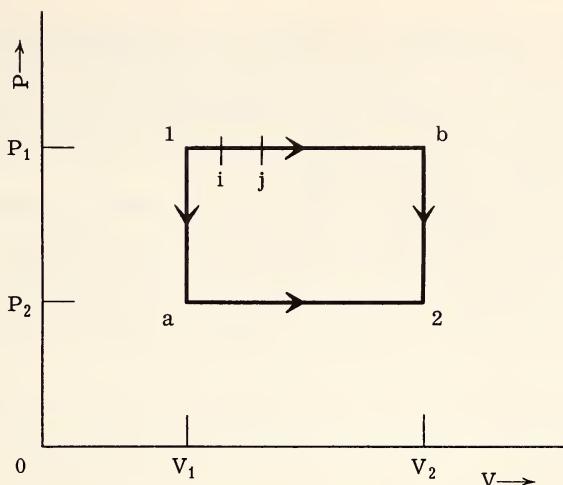


FIG. 4-1. Plots of quasistatic processes on the PV plane.

about the heat "contained" by the system, nor about the mechanical energy it contains. It possesses just one pool of contained energy, which we call its internal energy  $U$ , contributed to by input of *both* mechanical work and also of heat, which can be withdrawn either as mechanical energy or as heat. Any change in  $U$ ,  $dU$  is the difference between the heat added,  $dQ$ , and the work done by the system  $dW$  during a quasistatic process,

$$\begin{aligned}
 dU &= dQ - dW \\
 &= dQ - P dV + J dL + H dM + \mu dn + \dots
 \end{aligned} \tag{4-1}$$

where  $dU$  is a perfect differential and  $dQ$  and  $dW$  are imperfect ones. Note the convention used here;  $dQ$  is the heat *added to* the system,  $dW$  is the work *done by* the system.

This set of equations is the *first law of thermodynamics*. It states that mechanical work and heat are two forms of energy and must be lumped together when we compute the change in internal energy of the system. It was not obvious to physicists of the early nineteenth century. To have experiments show that heat could be changed into work *ad libitum*, that neither  $dQ$  nor  $dW$  were perfect differentials, seemed at the time to introduce confusion into a previously simple, symmetric theory.

There were some compensations. Gone was the troublesome question of how to measure the heat "contained" by the system. The question has no meaning; there is no state variable  $Q$ , there is only internal energy  $U$  to measure. Also the amount of heat  $dQ$  added could

sometimes be most easily measured by measuring  $dU$  and  $dW$  and computing  $dQ = dU + dW$ . An accurate measurement of the amount of heat added is even now difficult to make in many cases (the heat produced by passage of electric current through a resistance is relatively easy to measure, but the direct measurement of heat produced in a chemical reaction is still not easy).

Of course, compensations or not, Eq. (4-1) was the one that corresponded with experiment, so it was the one to use, and people had to persuade themselves that the new theory was really more "simple" and "obvious" than the old one. By now this revision of simplicity has been achieved; the idea of heat as a separate substance appears to us "illogical."

Just as with work, the total amount of heat added or withdrawn from a system depends on the process, on the path in the  $P, V$  plane of Fig. 4-1, for example. Of course the process must be that slow, stepwise kind, called quasistatic, if we are to use our thermodynamic formulas to calculate its change. To go from 1 to a in Fig. 4-1 we must remove enough heat from the gas, keeping its volume constant meanwhile, to lower its temperature from  $T_1 = (P_1 V_1 / nR)$  to  $T_a = (P_2 V_1 / nR)$ . We could do this relatively quickly (but not quasistatically) by placing the gas in thermal contact with a constant-temperature heat source at temperature  $T_a$ . Such a source, sometimes called a *heat reservoir*, is supposed to have such a large heat capacity that the amount of heat contributed by the gas will not change its temperature. In this case the gas would not be in thermal equilibrium until it settled down once more into equilibrium at  $T = T_a$ . To carry out a quasistatic process, for which we could use our formulas to compute the heat added, we should have to place the gas first into contact with a heat reservoir at temperature  $T_1 - dT$ , allowing it to come to equilibrium, then place it in contact with a reservoir at  $T_1 - 2dT$ , and so on.

To be sure, if the gas is a perfect gas of point atoms, we already know that  $U = (3/2)PV$ , so that  $U_a - U_1 = -(3/2)V_1(P_1 - P_2)$ , whether the system passes through intermediate equilibrium states or not, as long as states 1 and a are equilibrium states. Then since in this case  $dW = 0$ , we can immediately find  $dQ$ . But if we did not know the formula for  $U$ , but only knew the heat capacity of the gas at constant volume, we should be required (conceptually) to limit the process of going from 1 to a to a quasistatic one, in order to use  $C_V$  to compute the heat added. For a quasistatic process, for a perfect gas of point atoms where  $C_V = (3/2)nR$ , the heat added to the gas between 1 and a is

$$Q_{1a} = \int C_V dT = \left( \frac{3}{2} nR \right) \int (V_1 dP/nR) = -\frac{3}{2} V_1 (P_1 - P_2)$$

checking with the value calculated from the change in  $U$ .

### Quasistatic Processes

In going from a to 2 the same problem arises. We can imagine that the gas container is provided with a piston, which can be moved to change the volume  $V$  occupied by the gas. We could place the gas in thermal contact with a heat reservoir at temperature  $T_2 = (P_2 V_2 / nR)$  and also move the piston so the volume changes rapidly from  $V_1$  to  $V_2$  and then wait until the gas settles down to equilibrium. In this case we can be sure that the internal energy  $U$  will end up having the value  $(3/2)nRT_2 = (3/2)P_2 V_2$ , but we cannot use thermodynamics to compute how much work was done during the process or how much heat was absorbed from the heat reservoir. If, for example, instead of moving a piston, we turned a stopcock and let the gas expand freely into a previously evacuated volume  $(V_2 - V_1)$ , the gas would do no work while expanding. Whereas, if we moved the piston very slowly, useful work would be done and more heat would have to be taken from the reservoir in order to end up with the same value of  $U$  at state 2. In the case of free expansion, the energy not given up as useful work would go into turbulence and sound energy, which would then degenerate into heat and less would be taken from the reservoir by the time the system settled down to state 2.

If we did not know how  $U$  depends on  $P$  and  $T$ , but only knew the value of the heat capacity at constant pressure [which we shall show later equals  $(5/2)nR$  for a perfect gas of point atoms] we should have to devise a quasistatic process, going from a to 2, for which to compute  $\Delta Q_{a2}$  and  $\Delta W_{a2}$  and thence, by Eq. (4-1), to obtain  $\Delta U$ . For example, we can attach the piston to a device (such as a spring) which will maintain a constant pressure  $P_2$  on the gas no matter what position the piston takes up (such a device could be called a constant-pressure work source, or a *work reservoir*). We then place the gas in contact with a heat reservoir at temperature  $T_a + dT$ , wait until the gas comes to equilibrium at slightly greater volume, place it in contact with another reservoir at temperature  $T_a + 2dT$ , and so on. The work done in this quasistatic process at constant pressure is, as we said earlier,  $\Delta W_{a2} = P_2(V_2 - V_1)$ . The heat donated by the heat reservoir [if  $C_p = (5/2)nR$ ] is  $\Delta Q_{a2} = (5/2)nR \int dT = (5/2)P_2(V_2 - V_1)$  and the difference is  $\Delta U = \Delta Q_{a2} - \Delta W_{a2} = (3/2)P_2(V_2 - V_1)$ , as it must be.

Thus thermodynamic computations, using an appropriate quasistatic process, can predict the change in internal energy  $U$  (or in any other state variable) for any process, fast or slow, which begins and ends in an equilibrium state. But these calculations cannot predict the amount of intake of heat or the production of work during the process unless the process differs only slightly from the quasistatic one used in the calculations. It behooves us to avoid incomplete differentials, such as  $dW$  and  $dQ$ , and to express the thermodynamic changes in a system during a process in terms of state variables, which can

be computed for any equilibrium state, no matter how the system actually arrived at the state.

### Heat Capacities

To integrate  $U$  for a simple system, where

$$dU = dQ - P dV \quad (4-2)$$

we need to work out some relationships between the heat capacities and the partial derivatives of  $U$ . For example, if  $T$  and  $V$  are chosen to be the independent variables, the heat absorbed in a quasistatic process is

$$dQ = dU + P dV = (\partial U / \partial T)_V dT + [(\partial U / \partial V)_T + P] dV \quad (4-3)$$

Since  $C_V$  is defined as the heat absorbed per unit increase in  $T$ , when  $dV = 0$ , we see that

$$C_V = (\partial U / \partial T)_V \quad (4-4)$$

so that Eq. (4-3) can be written

$$dQ = C_V dT + [(\partial U / \partial V)_T + P] dV \quad (4-5)$$

If  $T$  and  $V$  are varied so that  $P$  remains constant, then when  $T$  changes by  $dT$ ,  $V$  will change by  $(\partial V / \partial T)_P dT$  and the amount of heat absorbed is

$$dQ = C_p dT = C_V dT + [(\partial U / \partial V)_T + P] (\partial V / \partial T)_P dT \quad \text{or}$$

or

$$C_p = C_V + (\partial V / \partial T)_P [(\partial U / \partial V)_T + P] \quad (4-6)$$

In our earlier discussion we stated that for a perfect gas of point atoms  $C_V = (3/2)nR$  and  $C_p = (5/2)nR$ ; we can now justify our statements. From Eq. (3-3) we know that for such a gas  $U = (3/2)nRT$ , so Eq. (4-4) gives us  $C_V$  immediately. It also shows that, for this gas  $(\partial U / \partial V)_T = 0$ , so that, from Eq. (4-6),

$$C_V = \frac{3}{2}nR; \quad C_p = C_V + \left( \frac{nRT}{V} \right) \frac{V}{T} = \frac{5}{2}nR \quad (4-7)$$

for a perfect gas of point atoms.

A similar set of relationships can be derived for other pairs of mechanical variables. For example, for paramagnetic materials, the specific heats for constant  $\mathfrak{M}$  and for constant  $\mathfrak{K}$  are obtained from Eq. (4-1) (assuming that  $V$ ,  $L$ , and  $n$  are constant):

$$dQ = (\partial U / \partial T)_{\mathfrak{M}} dT + [(\partial U / \partial \mathfrak{M})_T - \mathfrak{C}] d\mathfrak{M}$$

from which we can obtain

$$C_{\mathfrak{M}} = \left( \frac{\partial U}{\partial T} \right)_{\mathfrak{M}} ; \quad C_{\mathfrak{C}} = C_{\mathfrak{M}} - \left( \frac{\partial \mathfrak{M}}{\partial T} \right)_{\mathfrak{C}} \left[ \mathfrak{C} - \left( \frac{\partial U}{\partial \mathfrak{M}} \right)_T \right] \quad (4-8)$$

For a material obeying Curie's law  $\mathfrak{M} = (nD\mathfrak{C}/T)$ , it again turns out that  $(\partial U / \partial \mathfrak{M})_T = 0$ , analogous to the perfect gas, so that

$$C_{\mathfrak{C}} = C_{\mathfrak{M}} + (nD\mathfrak{C}^2 / T^2) = C_{\mathfrak{M}} + (1/nD)\mathfrak{M}^2 \quad (4-9)$$

Strictly speaking,  $C_{\mathfrak{M}}$  should be written  $C_{V\mathfrak{M}L} \dots$ , but since we are usually concerned with one pair of variables at a time, no ambiguity arises if we omit all but the variables of immediate interest.

### Isothermal and Adiabatic Processes

Other quasistatic processes can be devised beside those at constant volume and at constant pressure. For example, the system may be placed in thermal contact with a heat reservoir and the mechanical variables may be varied slowly enough so that the temperature of the system remains constant during the process. This is called an *isothermal* process. A heat capacity for this process does not exist (formally speaking,  $C_T$  is infinite). However it is important to be able to calculate the relationship between the heat  $dQ$  absorbed from the reservoir and the work  $dW$  done by the system while it proceeds.

For the perfect gas, where  $(\partial U / \partial V)_T = 0$ , and for paramagnetic materials, where  $(\partial U / \partial \mathfrak{M})_T = 0$ , and for other systems where  $U$  turns out to be a function of  $T$  alone, the heat absorbed from the reservoir during the isothermal process exactly equals the work done by the system. Such systems are perfect isothermal energy transformers, changing work into heat or vice versa without holding out any of it along the way. The transformation cannot continue indefinitely, however, for physical limits of volume or elastic breakdown or magnetic saturation or the like will intervene.

For less-simple substances the heat absorbed in an element of an isothermal process is

$$dQ = dU + dW = \left[ \left( \frac{\partial U}{\partial V} \right)_{T\mathfrak{M} \dots} + P \right] dV + \left[ \left( \frac{\partial U}{\partial \mathfrak{M}} \right)_{TV \dots} - H \right] d\mathfrak{M} \\ + \sum_i \left[ \left( \frac{\partial U}{\partial n_i} \right)_{TV \dots} - \mu_i \right] dn_i + \dots \quad (4-10)$$

differing from the work done by the amount by which  $U$  increases as

$V$  or  $M$  or  $n_i$  is changed isothermally. We remind ourselves that  $\mu_i dn_i$  is the chemical energy introduced into the system when  $dn_i$  moles of substance  $i$  is introduced or created in the system, and thus that  $-\mu_i dn_i$  is the chemical analogue of work done.

Another quasistatic process can be carried out with the system isolated thermally, so that  $dQ$  is zero. This is called an *adiabatic* process; for it the heat capacity of the system is zero. The relationship between the variables can be obtained from Eq. (4-1) by setting  $dQ = 0$ . For example, for a system with  $V$  and  $T$  as independent variables, using Eqs. (4-5) and (4-6), the change of  $T$  with  $V$  in an adiabatic process is

$$C_V dT = -\frac{C_p - C_v}{(\partial V / \partial T)_p} dV \quad \text{or} \quad \left(\frac{\partial T}{\partial V}\right)_s = -(\gamma - 1) \left(\frac{\partial T}{\partial V}\right)_p \quad (4-11)$$

where  $\gamma = (C_p/C_v)$  is a state variable. The reason for using the subscript  $s$  to denote an adiabatic process will be elucidated in Chapter 6. We see that when  $\gamma$  is constant (as it is for a perfect gas) the adiabatic change of  $T$  with  $V$  is proportional to the change of  $T$  with  $V$  at constant pressure.

For the perfect gas of point atoms, where  $C_p = (5/2)nR$  and  $C_v = (3/2)nR$ ,  $\gamma = 5/3$  and  $(\partial T / \partial V)_p = (P/nR) = (T/V)$ , the relation between  $T$  and  $V$  for an adiabatic expansion is

$$(dT/T) + (\gamma - 1)(dV/V) = 0 \quad \text{or} \quad TV^{\gamma-1} = (PV^{\gamma}/nR) = \text{const.} \quad (4-12)$$

Compressing a gas adiabatically increases its temperature, because  $\gamma > 1$  pressure increases more rapidly, with change of volume, in an adiabatic compression than in an isothermal compression, where  $(PV/nR)$  is constant.

Similarly, for a paramagnetic material that obeys Curie's law and happens to have  $C_{\mathfrak{M}}$  independent of  $T$  and  $\mathfrak{M}$ , the relation between  $T$  and  $\mathfrak{M}$  during adiabatic magnetization is [see Eqs. (4-8) and (4-9)]

$$C_{\mathfrak{M}} dT = H d\mathfrak{M} \quad \text{or} \quad C_{\mathfrak{M}} dT = \frac{\mathfrak{M}T}{nD} d\mathfrak{M} \quad \text{or}$$

$$T = T_0 \exp(\mathfrak{M}^2 / 2nDC_{\mathfrak{M}}) \quad (4-13)$$

When  $\mathfrak{M} = 0$ , the atomic magnets, responsible for the paramagnetic properties of the material, are rotating at random with thermal motion; impressing a magnetic field on the material tends to line up the magnets and reduce their thermal motion and so to "squeeze out"

their heat energy, which must go into translational energy of the atoms (increased temperature) since heat is not removed in an adiabatic process. Reciprocally, if a paramagnetic material is magnetized, brought down to as low a temperature as possible and then demagnetized, the material's temperature will be still further reduced. By this process of adiabatic demagnetization, paramagnetic materials have been cooled from about  $1^{\circ}\text{K}$  to less than  $0.01^{\circ}\text{K}$ , the closest to absolute zero that has been attained.

# 5

## The Second Law of Thermodynamics

Once it had been demonstrated that heat is a form of energy, the proof of the first law of thermodynamics became merged with the proof of the conservation of energy. The experimental authentication of the second law is less direct; in a sense it is evidenced by the success of thermodynamics as a whole. The enunciation of the second law is also roundabout; its various paraphrases are many and, at first sight, unconnected logically. We could introduce the subject by asking whether there exists an integration factor for the imperfect differential  $dQ$ , or by asking for a quantitative measure of the difference between the quasistatic process leading from 1 to a in Fig. 4-1 and the more rapid process of placing the gas immediately in contact with a heat reservoir at temperature  $T_a$ , or else by asking whether 100 per cent of the heat withdrawn from a heat reservoir can be converted into mechanical work, no matter how much is withdrawn. We shall start with the last question and we find that in answering it we answer the other two.

### Heat Engines

Chemical or nuclear combustion can provide a rough equivalent of a constant temperature heat reservoir; as heat is withdrawn more can be provided by burning more fuel. Can we arrange it so this continuous output of heat energy is all continuously converted into mechanical work? The second law of thermodynamics answers this question in the negative, and provides a method of computing the maximum fraction of the heat output which *can* be changed into work in various circumstances.

At first sight this appears to contradict a result obtained in Chapter 4. There it was pointed out that a system, with internal energy that is a function of temperature only (such as a perfect gas or a perfect paramagnetic material), when placed in contact with a constant-temperature heat source, can isothermally transform all the heat it withdraws from the reservoir into useful work, either mechanical or

electromagnetic. The trouble with such a process is that it cannot continue to do this *indefinitely*. Sooner or later the pressure gets too low or the tension gets greater than the elastic limit or the magnetic material becomes saturated, and the transformer's efficiency drops to zero. What is needed is a *heat engine*, a thermodynamic system that can operate cyclically, renewing its properties periodically, so it can continue to transform heat into work indefinitely.

Such an engine cannot be built to run entirely at one temperature, that of the heat source. If it did so the process would be entirely isothermal, and if we try to make an isothermal process cyclic by reversing its motion (compressing the gas again, for example) we find we are taking back all the work that has been done and reconverting it into heat; returning to the start leaves us with no net work done and all the heat given back to the reservoir. Our cycle, to result in net work done and thus net heat withdrawn, must have some part of it operating at a lower temperature than that of the source. And thus we are led to the class of cyclical operations called Carnot cycles.

### Carnot Cycles

A Carnot cycle operates between two temperatures, a hotter,  $T_h$ , that of the heat source, and a colder,  $T_c$ , that of the heat sink. Any sort of material can be used, not just one having  $U$  a function of  $T$  only. And any pair of mechanical variables can be involved,  $P$  and  $V$  or  $J$  and  $L$  or  $\pi$  and  $\mathfrak{M}$  (we shall use  $P$  and  $V$  just to make the discussion specific). The cycle consists of four quasistatic operations: an isothermal expansion from 1 to 2 (see Fig. 5-1) at temperature  $T_h$ ,

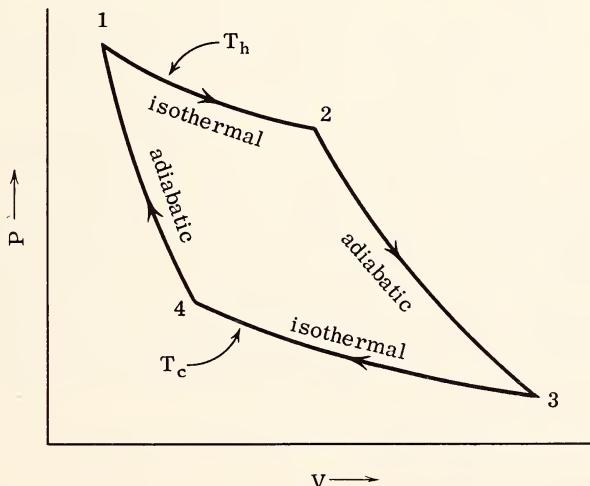


FIG. 5-1. Example of a Carnot cycle, plotted in the PV plane.

withdrawing heat  $\Delta Q_{12}$  from the source and doing work  $\Delta W_{12}$  (not necessarily equal to  $\Delta Q_{12}$ ); an adiabatic expansion from 2 to 3, doing further work  $\Delta W_{12}$  but with no change in heat, and ending up at temperature  $T_C$ ; an isothermal compression at  $T_C$  from 3 to 4 requiring work  $-\Delta W_{34} = \Delta W_{43}$  to be done on the system and contributing heat  $-\Delta Q_{34} = \Delta Q_{43}$  to the heat sink at temperature  $T_C$ , ending at state 4, so placed that process 4 to 1 can be an adiabatic compression, requiring work  $-\Delta W_{41} = \Delta W_{14}$  ( $\Delta Q_{41} = 0$ ) to be done on the system to bring it back to state 1, ready for another cycle (Fig. 5-1). This is a specialized sort of cycle but it is a natural one to study and one that in principle should be fairly efficient. Since the assumed heat source is at constant temperature, part of the cycle had better be isothermal, and if we must "dump" heat at a lower temperature, we might as well give it all to the lowest temperature reservoir we can find. The changes in temperature should thus be done adiabatically.

This cycle, of course, does not convert all the heat withdrawn from the reservoir at  $T_h$  into work; some of it is dumped as unused heat into the sink at  $T_C$ . The net work done by the engine per cycle is the area inside the figure 1234 in Fig. 5-1, which is equal to  $\Delta W_{12} + \Delta W_{23} + \Delta W_{34} + \Delta W_{41} = \Delta W_{12} + \Delta W_{23} - \Delta W_{43} - \Delta W_{14}$  and which, according to the first law, is equal to  $\Delta Q_{12} + \Delta Q_{34} = \Delta Q_{12} - \Delta Q_{43}$ . The efficiency  $\eta$  with which the heat withdrawn from the source at  $T_h$  is converted into work is equal to the ratio between the work produced and the heat withdrawn.

$$\eta = \frac{\Delta W_{12} + \Delta W_{23} - \Delta W_{43} - \Delta W_{14}}{\Delta Q_{12}} = \frac{\Delta Q_{12} - \Delta Q_{43}}{\Delta Q_{12}} = 1 - \frac{\Delta Q_{43}}{\Delta Q_{12}} \quad (5-1)$$

We note that, since all the operations are quasistatic, the cycle is *reversible*; it can be run backward, withdrawing heat  $\Delta Q_{43}$  from the reservoir at temperature  $T_C$  and depositing heat  $\Delta Q_{12}$  in the reservoir at  $T_h$ , requiring work  $\Delta Q_{12} - \Delta Q_{43}$  to make it go.

There are a large number of Carnot cycles, all operating between  $T_h$  and  $T_C$ ; ones using  $P$  and  $V$  to generate work, involving different substances with different equations of state; ones using  $\mathcal{C}$  and  $\mathcal{M}$  to produce magnetic energy, using different paramagnetic substances; and so on. One way of stating the second law is to say that all Carnot cycles operating between the temperatures  $T_h$  and  $T_C$  *have the same efficiency*. Another way is to say that *no engine*, or combination of engines, operating between a maximum temperature  $T_h$  and a minimum temperature  $T_C$  *can be more efficient* than any Carnot cycle operating between these temperatures.

### Statements of the Second Law

To show that these statements are equivalent we shall show that *if* we could find a cycle of greater efficiency than a Carnot cycle, oper-

ating between  $T_h$  and  $T_c$ , we can combine them to obtain a perfectly efficient engine, thus more efficient than either. Figure 5-2 shows how this can be done. We assume that the less efficient one is the

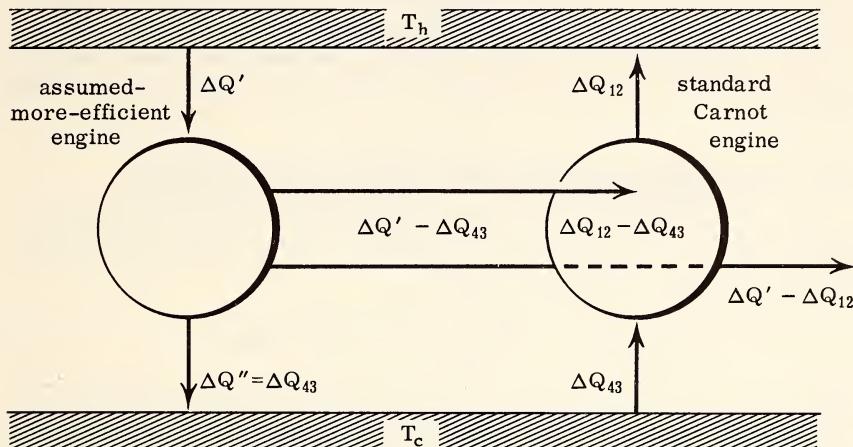


FIG. 5-2. Carnot engine (reversed) driven by an engine assumed more efficient; the combination would make a perfect engine, which is impossible.

“standard” one described in Eq. (5-1). We shall run this backward, requiring net work  $\Delta Q_{12} - \Delta Q_{43}$  to take heat  $\Delta Q_{43}$  from the lower temperature and depositing heat  $\Delta Q_{12}$  at the upper. We adjust the presumed better engine so its exhaust  $\Delta Q''$  at  $T_c$  is equal to  $\Delta Q_{43}$ , the same amount that the first engine withdraws. The amount of heat  $\Delta Q'$  it withdraws at  $T_h$  must be larger than  $\Delta Q_{12}$  if its efficiency  $1 - (\Delta Q''/\Delta Q')$  is to be larger than the value  $1 - (\Delta Q_{43}/\Delta Q_{12})$  for the “standard” engine. We now use this better engine to run the “standard” one in its reversed cycle. Actually there will be work left over, an amount  $(\Delta Q' - \Delta Q_{43}) - (\Delta Q_{12} - \Delta Q_{43}) = \Delta Q' - \Delta Q_{12}$ , which can be used as we please. The combined engine thus withdraws a net heat  $\Delta Q' - \Delta Q_{12}$  from the upper heat reservoir, dumps no net heat into the lower, and produces net work  $\Delta Q' - \Delta Q_{12}$ ; it is a perfect engine. Thus a contradiction of the first statement above leads to a contradiction of the second statement. If there can be no engine more efficient than a Carnot cycle, then all Carnot cycles (between  $T_h$  and  $T_c$ ) must have the same efficiency.

We can see now that still another way of stating the second law is as follows: It is impossible to convert, continuously, heat from a reservoir at one temperature  $T_h$  into work, without at the same time transferring additional heat from  $T_h$  to a colder temperature  $T_c$ . This way of stating it is called *Kelvin's principle*. Still another way

is to state that it is impossible to transfer, in a continuous manner, heat from a lower temperature reservoir to one at higher temperature without at the same time doing work to effect the transfer. This is called *Clausius' principle*. Its equivalence to the other three statements can be demonstrated by manipulating the combination of Fig. 5-2; by running it backward and adjusting the assumed better engine so that  $\Delta Q' - \Delta Q'' = \Delta Q_{12} - \Delta Q_{43}$  and thus  $\Delta Q'' < \Delta Q_{43}$  and  $\Delta Q' > \Delta Q_{12}$ , for example.

### The Thermodynamic Temperature Scale

If all Carnot cycles operating between the same pair of temperatures,  $T_h$  and  $T_c$ , have the same efficiency, this efficiency must be simply a function of  $T_h$  and  $T_c$ :

$$\eta = 1 - \Psi(T_h, T_c); \quad \Psi(T_h, T_c) = (\Delta Q_{43} / \Delta Q_{12}) \quad (5-2)$$

The ratio of the heat dumped to that withdrawn must be the same for all these cycles. To find how  $\Psi$  depends on the temperatures  $T_h$  and  $T_c$ , as measured by a thermometer, we break up a Carnot cycle into two cycles, each using the same material, as shown in Fig. 5-3. The upper one takes heat  $\Delta Q_{12}$  from the upper reservoir at a temperature  $\theta_h$  (on the scale of the thermometer we are using) does work  $\Delta Q_{12} - \Delta Q_{65}$ , and delivers heat  $\Delta Q_{65}$  to an intermediate reservoir at a measured temperature  $\theta_m$ . This reservoir immediately passes on this heat  $\Delta Q_{65}$  to the second engine, which produces work  $\Delta Q_{65} - \Delta Q_{43}$  and delivers heat  $\Delta Q_{43}$  to a reservoir at temperature  $\theta_c$  as measured on our thermometer. The combination, which produces a total work of  $\Delta Q_{12} - \Delta Q_{43}$ , is thus completely equivalent to a single Carnot cycle, using the same material and operating between  $\theta_h$  and  $\theta_c$  on our scale, withdrawing  $\Delta Q_{12}$  from the upper, exhausting  $\Delta Q_{43}$  at the lower, and doing work  $\Delta Q_{12} - \Delta Q_{43}$ . Therefore, according to Eq. (5-2), the efficiencies  $\eta_u$  and  $\eta_l$  of the two component cycles and the efficiency  $\eta_c$  of the combination, considered as a single engine, are related as follows:

$$\begin{aligned} \eta_u &= 1 - \Psi(\theta_h, \theta_m); \quad \Psi(\theta_h, \theta_m) = \frac{\Delta Q_{65}}{\Delta Q_{12}} \\ 1 - \eta_l &= \Psi(\theta_m, \theta_c) = \frac{\Delta Q_{43}}{\Delta Q_{65}} \\ 1 - \eta_c &= \Psi(\theta_h, \theta_c) = \frac{\Delta Q_{43}}{\Delta Q_{12}} = \Psi(\theta_h, \theta_m)\Psi(\theta_m, \theta_c) \end{aligned} \quad (5-3)$$

For the equation relating the three values of the function  $\Psi$ , for the three pairs of values of measured temperature  $\theta$ , to be valid,  $\Psi$

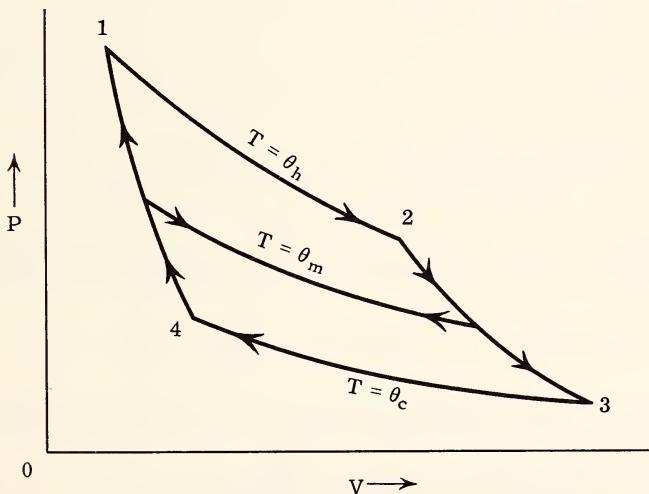
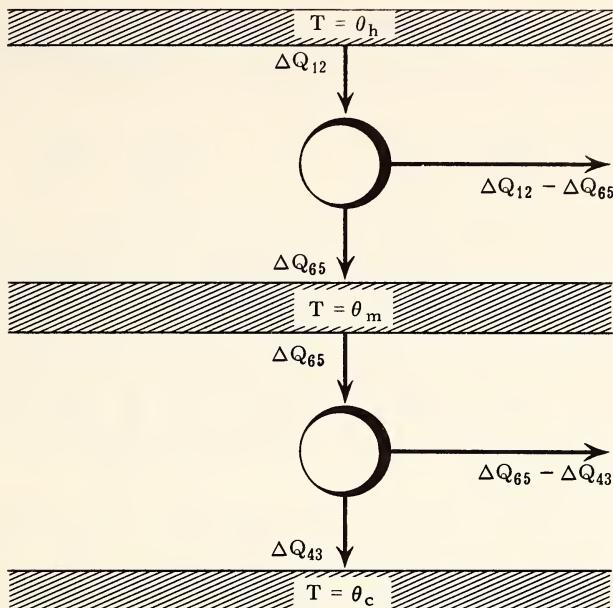


FIG. 5-3. Arrangement of two Carnot cycles so their combined effect is equivalent to one cycle between the temperature extremes.

must have the functional form  $\Psi(x,y) = [T(y)/T(x)]$ , where  $T(\theta)$  is some single-valued, monotonically increasing function of  $\theta$ , the temperature reading on the thermometer used. For then  $\Psi(\theta_m, \theta_c)\Psi(\theta_h, \theta_m)$  will equal

$$[T(\theta_C)/T(\theta_M)][T(\theta_M)/T(\theta_h)] = [T(\theta_C)/T(\theta_h)] = \Psi(\theta_h, \theta_C)$$

Therefore,

$$\Psi(\theta_h, \theta_C) = \frac{\Delta Q_{43}}{\Delta Q_{12}} = \frac{T(\theta_C)}{T(\theta_h)} \quad \text{or} \quad \frac{\Delta Q_{43}}{T(\theta_C)} = \frac{\Delta Q_{12}}{T(\theta_h)} \quad (5-4)$$

We can thus experimentally determine the function  $T(\theta)$  by using various Carnot cycles, all having the same upper temperature, read as  $\theta_h$  on our thermometer, but each going to a different lower temperature, discarding a different amount of heat and thus doing different amounts of work. By measuring the common value of  $\Delta Q_{12}$  and the different values of the heat discarded,  $\Delta Q_{43}$ , we can compute  $T(\theta_C)$  as  $(\Delta Q_{43}/\Delta Q_{12})$  times the common value  $T(\theta_h)$ . If, for example, a cycle with lower reading  $\theta_d$  has its discarded heat just 1/2 of  $\Delta Q_{12}$ , then  $T(\theta_d) = (1/2)T(\theta_h)$ .

The numerical values of  $\theta_h$  and all the  $\theta_C$ 's were obtained from the arbitrary scale of the particular thermometer used. It would seem more appropriate to use  $T(\theta)$  itself, rather than  $\theta$ , for a temperature scale. We can use, for our upper reservoir, the temperature of melting ice, and set  $T(\theta_h) = 273^\circ$ . The value of  $T(\theta_C)$  for any colder temperature (such as the boiling point of oxygen, for example) is then  $(\Delta Q_{43}/\Delta Q_{12}) \times 273^\circ$ , where  $\Delta Q_{12}$  and  $\Delta Q_{43}$  are the heats involved in the Carnot cycle operating between the melting point of ice and the boiling point of oxygen. Such a scale of temperature, as determined by measured heat ratios for Carnot cycles, is called the *thermodynamic scale*, and measurements given in this scale are in degrees Kelvin.

This now completes our series of definitions of temperature started in Chapter 1. From now on temperature  $T$  will always be measured in degrees Kelvin. In its terms, the efficiency of a Carnot cycle operating between  $T_h$  and  $T_C$  (both measured in degrees Kelvin) is

$$\eta = 1 - \Psi(T_h, T_C); \quad \Psi(T_h, T_C) = \frac{\Delta Q_{43}}{\Delta Q_{12}} = \frac{T_C}{T_h} \quad (5-5)$$

This is the maximum efficiency we can get from an engine that operates between  $T_h$  and  $T_C$ .

Thus the second law is a sort of relativistic principle. The minimal temperature at which we can exhaust heat is determined by the temperature of our surroundings, and this limits the efficiency of transfer of heat into work. Heat at temperatures high compared to our surroundings is "high-quality" heat; if we handle it properly a large portion of it can be changed into useful work. Heat at temperature twice that of our surroundings (on the Kelvin scale) is already half degraded; only half of it can be usefully employed. And heat at the temperature of our surroundings is useless to us for getting work done. Even heat at

a million degrees Kelvin would be useless if the whole universe were at this same temperature. Temperature *differences* enable us to produce mechanical energy, not absolute magnitudes of average temperature.

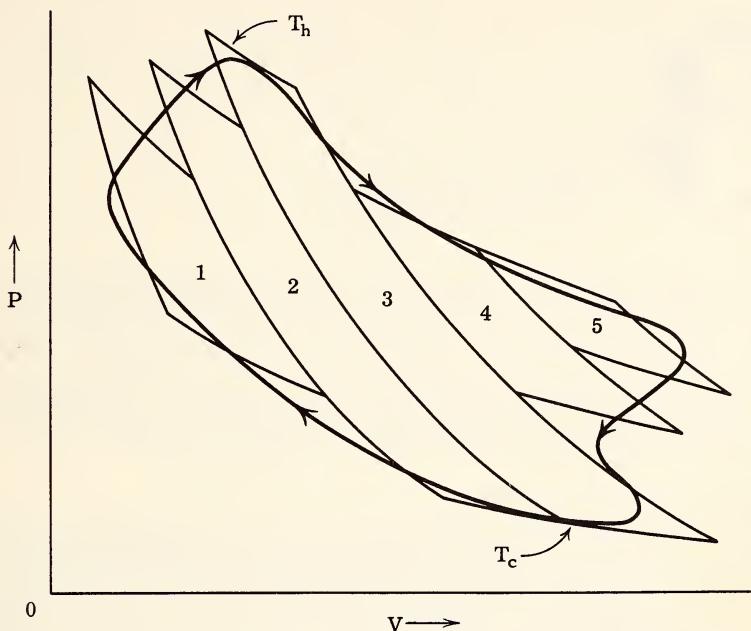


FIG. 5-4. Reversible cycle (heavy line) simulated by a combination of several Carnot cycles.

In principle we can build up a combination of Carnot cycles to simulate any kind of reversible cycle, such as the one shown by the heavy line in Fig. 5-4. In such a cycle, heat is taken on and given off at different temperatures, none of the elementary processes being isothermal or adiabatic. The maximum temperature reached is  $T_h$ , for the isothermal curve tangent to the top of the loop, and the minimum is  $T_c$ , for the lower tangent isothermal. The work produced is the area within the heavy line. This cycle is crudely approximated by the five Carnot cycles shown, with their isothermals and adiabatics as light lines; a better approximation could be obtained with a large number of Carnot cycles. The efficiency of subcycle 3 is greatest, because it operates between the greatest spread of temperatures; the others have less efficiency. Thus any cycle that takes in or gives off heat while the temperature is changing is not as efficient as a Carnot cycle operating between the same maximum and minimum temperatures, i.e., which takes on all its heat at  $T_h$  and which gives up all its heat at  $T_c$ .

# 6

## Entropy

We notice that, for a Carnot cycle, the relationship between each element  $dQ$  of heat taken on and the thermodynamic temperature  $T$  at which it is taken on (or given off) is such that the integral of  $dQ/T$  taken completely around the cycle is zero. The heat taken on at  $T_h$  is  $\Delta Q_{12}$  and the heat "taken on" at  $T_c$  is the negative quantity  $\Delta Q_{34} = -\Delta Q_{43}$ ; Eq. (5-5) states that the sum  $(\Delta Q_{12}/T_h) - (\Delta Q_{43}/T_c) = 0$ .

### A Thermal-State Variable

Since any quasistatic, reversible cycle can be considered as a sum of Carnot cycles, as in Fig. 5-4, we see that for any such cycle the integral of the quantity  $dQ/T$  around the whole cycle is zero. But for any thermodynamic state function  $Z(x,y)$  (as in Fig. 6-1) the integral of the perfect differential  $dZ$  around a closed path (such as ABA in Fig. 6-1) is zero, as long as all parts of the path are reversible processes; alternately any differential that integrates to zero around any closed path is a perfect differential and its integral is a state function of the variables  $x,y$ .

Therefore the quantity  $dS = dQ/T$  is a perfect differential, where  $dQ$  is the heat given to the system in an elementary, reversible process and  $T$  is the thermodynamic temperature of the system during the process. The integral of this perfect differential,  $S(x,y)$ , is a state variable and is called the *entropy* of the system. It is an extensive variable, proportional to  $n$ .

This result, which is still another way of stating the second law, can be rephrased to answer the first of the questions posed in the first paragraph of Chapter 5. There *is* an integrating factor for  $dQ$ , if heat  $dQ$  is absorbed in a reversible process; it is the reciprocal of the thermodynamic temperature, defined in Eq. (5-5). The resulting perfect differential  $dQ/T$  measures the change  $dS$  in the state variable  $S$ , the entropy; and the difference  $S_2 - S_1$  of entropy between equilibrium states 1 and 2 is computed by integrating  $dQ/T$  along any re-

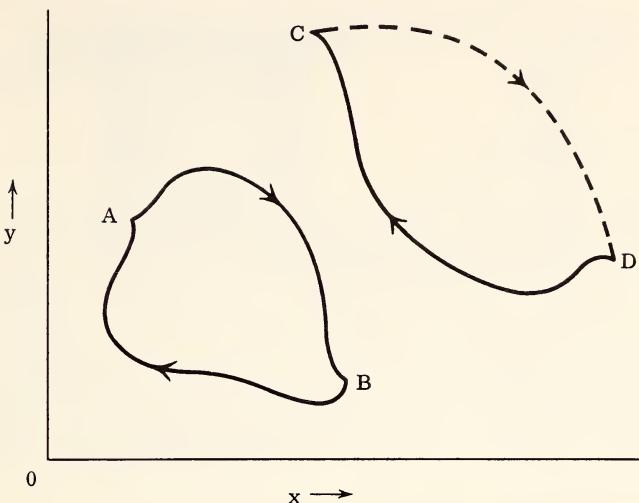


FIG. 6-1. Paths in the  $xy$  plane for reversible processes (solid lines) and for spontaneous processes (dashed lines).

versible path between 1 and 2. On the other hand, there is no integrating factor for  $dQ$  for an irreversible process.

The entropy  $S$  is the extensive variable that pairs with  $T$  as  $V$  does with  $P$  and  $\mathfrak{M}$  with  $\mathfrak{C}$ . The heat taken on by the system in a reversible process is  $dQ = T dS$ , just as the work done by the system is  $dW = P dV$  or  $-\mathfrak{C} d\mathfrak{M}$ . Thus the equation that represents *both the first and second laws of thermodynamics* is

$$dU = T dS - P dV + J dL + \mathfrak{C} d\mathfrak{M} + \sum_i \mu_i dn_i + \dots \quad (6-1)$$

for a reversible process. This equation, plus the empirical heat expressions for heat capacity and the equations of state, constitutes the mathematical model for thermodynamics, by means of which we can predict the thermal behavior of substance in bulk.

Basic equation (6-1) can be integrated to obtain  $U$ , once we know the dependence of  $T$ ,  $P$ ,  $J$ , etc., on  $S$ ,  $V$ ,  $L$ , etc., which we can take to be the independent variables of the system. Of course we can, if we choose, use  $P$  instead of  $V$  as one of the independent variables but, if we desire to calculate  $U$ , Eq. (6-1) shows that the extensive variables  $S$ ,  $V$ ,  $L$ , etc., are the "natural" ones to use. When expressed as function of the extensive variables,  $U$  has the properties of a potential function, for its partial with respect to one of the extensive variables ( $V$ , for example) is equal to the corresponding thermodynamic "force," the related intensive function ( $P$ , for example). Thus

$$T = \left( \frac{\partial U}{\partial S} \right)_{V, \mathfrak{M}, \dots} ; \quad P = - \left( \frac{\partial U}{\partial V} \right)_{T, \mathfrak{M}, \dots} ; \quad \mathcal{E} = \left( \frac{\partial U}{\partial \mathfrak{M}} \right)_{T, V, \dots} ;$$

$$\mu_i = \left( \frac{\partial U}{\partial n_i} \right)_{T, V, \dots} \quad (6-2)$$

Because Eq. (6-1) is a sum of intensive variables, each multiplied by the differential of its corresponding extensive variable, we can apply a trick devised by Euler to calculate  $U$ . Let the  $\nu$  extensive variables appropriate for a system be  $X_1, \dots, X_\nu$ ; then the corresponding intensive variables are  $Y_j = \partial U / \partial X_j$ , all of which are uniform throughout the system at equilibrium. Now we increase the amount of material by a factor  $\lambda$ , keeping all the intensive variables  $Y_j$  constant during the change. The internal energy  $U$  for the new system is just  $\lambda$  times the  $U$  of the original system, and the extensive variables are also increased by  $\lambda$ , so that

$$U(\lambda X_1, \dots, \lambda X_\nu) = \lambda U(X_1, \dots, X_\nu)$$

Differentiating with respect to  $\lambda$  on both sides and using the definitions of the intensive variables as partials of  $U$ , we have

$$\frac{d}{d\lambda} U(\lambda X_1, \dots, \lambda X_\nu) = \sum_{j=1}^{\nu} (\partial U / \partial X_j) \lambda X_j = \sum_{j=1}^{\nu} Y_j X_j = U(X_1, \dots, X_\nu)$$

Thus, in terms of our familiar variables,

$$U(S, V, L, \mathfrak{M}, n_1, n_2, \dots) = ST - PV + JL + \mathcal{E}\mathfrak{M} + \sum_i \mu_i n_i + \dots \quad (6-3)$$

This is *Euler's equation*, which will be of use later. It may be considered to be the basic equation of thermodynamics; all the rest may be derived from it.

### Reversible Processes

We are now in a position to be more specific about the adjective *reversible*, which we first used for a cycle (such as a Carnot cycle) and which we recently have been applying to processes. To see what it means let us first consider a few *irreversible* processes. Suppose a gas is confined at pressure  $P_0$  within a volume  $V_0$  of a thermally insulated enclosure, as shown in Fig. 6-2. The gas is confined to  $V_0$  by a diaphragm  $D$ ; the rest of the volume,  $V_1 - V_0$ , is evacuated. We then break the diaphragm and let the gas undergo free expansion until it comes to a new equilibrium at volume  $V_1$ . This is a *spontaneous* process, going automatically in one direction only. It is obviously irreversible; the gas would never return by itself to volume  $V_0$ .

Next suppose we place an object, originally at temperature  $T_h$ , in thermal contact with a heat reservoir at temperature  $T_c$ , less than

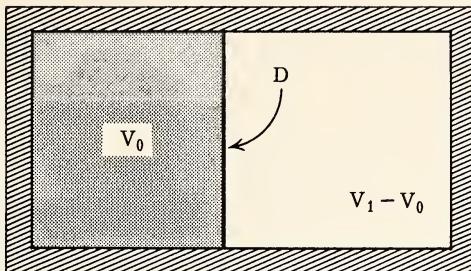


FIG. 6-2. The Joule experiment.

$T_h$ . Here again the process is spontaneous; heat flows from the object until it comes to equilibrium at temperature  $T_c$ . This also is an irreversible process; it would take work (or heat from a reservoir at  $T_h$ ) to warm the body up again.

We can thus define the adjective "reversible" in a negative way; a reversible process is one that has no irreversible portion. To expand the gas from  $V_0$  to  $V_1$  reversibly we could replace the diaphragm by a piston and move it slowly to the right. During the motion, as the volume is increased by  $dV$ , the gas is never far from equilibrium, and a reversal of motion of the piston (so the volume decreases again by  $dV$ ) would bring the gas back to its earlier state. In such a case we could retrace every part of the process in detail. Every reversible process is quasistatic; not all quasistatic processes are reversible.

In Eq. (6-1) we pointed out that the integral of  $dQ/T$  around a reversible cycle is zero. If the cycle is irreversible the integral differs from zero; the second law requires it to be *less than zero*. For example, suppose the irreversible cycle took in all its heat, an amount  $\Delta Q'$  at  $T_h$  and exhausted an amount  $\Delta Q''$  all at  $T_c$ . The efficiency of such a cycle would have to be less than the value  $1 - (T_c/T_h) = 1 - (\Delta Q_{43}/\Delta Q_{12})$  for a Carnot cycle between the same temperatures. This means that  $\Delta Q'$  would have to be smaller than  $\Delta Q_{12}$  or  $\Delta Q''$  would have to be larger than  $\Delta Q_{43}$ , or both, so that the integral of  $dQ/T$  around the irreversible cycle would turn out less than that for the Carnot cycle, i.e., less than zero. The argument can be generalized for all closed cycles. Thus another way of stating the second law is that for all closed cycles the integral around the cycle,

$$\oint (dQ/T) \leq 0 \quad (6-4)$$

where the equality holds for reversible cycles and the inequality is for irreversible ones. Since  $dS$  is measured by the value of  $dQ/T$  when the process is reversible, we also have that

$$dS \geq dQ/T \quad (6-5)$$

where again the equality holds for reversible processes, inequality for irreversible ones.

### Irreversible Processes

For example, suppose the dotted line from C at D in Fig. 6-1 represents a spontaneous process, during which no heat is absorbed or given up (as is the case with the free expansion of the gas in Fig. 6-2), so that  $\int (\delta Q/T) = 0$  for the dotted line. Since the process is irreversible,  $\int dS = S_D - S_C$  must be larger than  $\int (\delta Q/T) = 0$ . In other words, *during a spontaneous process taking place in a thermally isolated system, the entropy always increases*.

The statement is not so simple when the spontaneous process involves transfer of heat, as with the irreversible cooling of a body from  $T_h$  to  $T_c$  mentioned above [see also the discussion four paragraphs below Eq. (4-1)]. In this case the body loses entropy and the reservoir gains it. If the heat capacity of the body is  $C_v$ , a constant, the heat reservoir gains  $C_v(T_h - T_c)$  and, since the reservoir is always at  $T_c$ , its gain in entropy is  $C_v[(T_h - T_c)/T_c]$ . The loss in entropy of the body is not much harder to compute. During its spontaneous discharge of heat to the reservoir, and before it comes to uniform temperature  $T_c$ , the body is not in equilibrium, so  $dS$  does not equal  $dQ/T$ . However we can devise a quasistatic process, placing a poor heat conductor between the body and the reservoir, so the heat flows into the reservoir slowly and the body, at any time, will have a nearly uniform temperature  $T$ , where  $T$  starts at  $T_h$  and gradually drops to  $T_c$  at the end. The loss of entropy from the body is thus the integral of  $C_v(dT/T)$ , which is  $C_v \ln(T_h/T_c)$  if  $C_v$  is constant. This is always smaller than  $C_v[(T_h - T_c)/T_c]$ , the entropy gained by the reservoir, although the two approach each other in value as  $T_h$  approaches  $T_c$ . Thus, although the entropy of the body decreases during the spontaneous cooling of the body, the total entropy of body and reservoir (which we might call the entropy of the universe) increases by the amount

$$S = C_v x - C_v \ln(1+x) = C_v \left( \frac{1}{2} x^2 - \frac{1}{3} x^3 + \frac{1}{4} x^4 - \dots \right)$$

where  $x = [(T_h - T_c)/T_c]$ , which is positive for all values of  $x > -1$ .

Thus the statement at the end of the previous paragraph can be generalized by saying that in a spontaneous process of any kind, even if the entropy of the body decreases, the entropy of some other system increases even more, so that the *entropy of the universe always increases during an irreversible process*. This, finally, is the answer to the second question at the beginning of Chapter 5; the measure of the difference between a reversible and an irreversible process lies in the entropy change of the universe.

Entropy is a measure of the unavailability of heat energy. The entropy of a certain amount of heat at low temperature is greater than it is at high temperature, loosely speaking. Alternately, entropy measures the degree of disorganization of the system. Irreversible processes increase disorder, increase the amount of low-temperature heat, and thus increase the entropy of the universe. Reversible processes, on the other hand, simply transfer entropy from one body to another, keeping the entropy of the universe constant. A few examples will familiarize us with these ideas.

### Entropy of a Perfect Gas

The entropy of  $n$  moles of a perfect gas of point atoms, for which  $U = (3/2)nRT$  and  $PV = nRT$ , may be determined by integration of Eq. (6-1), which is for this case

$$T \, dS = dU + P \, dV \quad \text{or} \quad dS = \frac{3}{2}(nR/T) \, dT + (nR/V) \, dV$$

so

$$S = nR \ln \left[ (T/T_0)^{3/2} (V/V_0) \right] + S_0$$

where  $S = S_0$  when  $T = T_0$  and  $V = V_0$ . Increase in either  $T$  or  $V$  increases the entropy of the gas. Instead of  $T$  and  $V$  (and  $n$ ),  $S$  and  $V$  (and  $n$ ) can be used as independent variables, in which case

$$T = T_0 \left( \frac{V_0}{V} \right)^{2/3} e^{2(S - S_0)/3nR}; \quad P = \frac{nRT_0}{V_0} \left( \frac{V_0}{V} \right)^{5/3} e^{2(S - S_0)/3nR}$$

These formulas immediately provide us with the dependence of  $T$  and  $P$  on  $V$  for an adiabatic process. For a *reversible* adiabatic process,  $dS = dQ/T = 0$ , so  $S$  is constant, which is why partials for such processes are labeled with subscript  $S$ .

We can also use Euler's equation (6-3) to calculate the chemical potential per mole of point atoms. For this system,  $U = TS - PV + \mu n$ , the atoms being all of one kind. Inserting the expressions for  $U$ ,  $PV$ , and  $TS$  in terms of  $T$ ,  $V$ , and  $n$  and dividing by  $n$ , we find that

$$\mu = -Ts_0 + RT \ln [e^{5/2} (V_0/V)(T_0/T)^{3/2}] \quad (6-6)$$

where  $e$  is the base of the natural logarithms ( $\ln e = 1$ ) and  $s_0 = (S_0/n_0)$  is the entropy per mole at  $T_0$ ,  $V_0$ , and  $n_0$ . Therefore we can use  $\mu$  for an independent variable and obtain, for example,

$$V = V_0(T_0/T)^{3/2} \exp \left( -\frac{5}{2} - \frac{s_0}{R} - \frac{\mu}{RT} \right)$$

The function  $U(S, V, n)$  is  $(nb_0/V^{2/3}) e^{2S/3nR}$ , where constant  $b_0$  equals  $T_0(V_0/n_0)^{2/3} e^{-2S_0/3R}$ , independent of  $S$ ,  $V$ , and  $n$ .

### The Joule Experiment

Let us return, for a page or two, to the irreversible, free-expansion process pictured in Fig. 6-2. Initially the gas is confined to volume  $V_0$  and is at temperature  $T_0$ ; after the diaphragm has broken, the gas expands spontaneously and eventually settles down to equilibrium in volume  $V_1$  at a new temperature  $T_1$ . We should like to be able to compute  $T_1$  and also to calculate the increase in entropy during the process for any sort of gas, not just for a perfect gas. What we need to know to do this is the nature of the final equilibrium state; we then can use the appropriate version of Eq. (6-1),

$$dU = T dS - P dV \quad (6-7)$$

to integrate from initial to final state via a reversible path.

Equation (6-1), in its various forms, can represent *any* reversible process, therefore can be used to compute the difference of value of any state variable between any initial and final states. In the free-expansion case now under study, the actual process is far from reversible, but it starts and stops with equilibrium states and we can compute the difference between start and finish without bothering to learn what the system did in-between.

In the free-expansion case, since the system is thermally insulated during the process and no work is done, the internal energy must be the same at the finish as it was at the start. Thus the change of value of any state variable caused by free expansion may be computed by integrating its rate of change with respect to  $V$ , at constant  $U$ . How this partial may be found, by using Eqs. (3-6) and (3-7) to manipulate Eq. (6-7), goes as follows.

First we express  $dS$  and  $dU$  in terms of their partials and the differentials of the independent variables  $T$  and  $V$  and equate coefficients of these differentials:

$$T \left( \frac{\partial S}{\partial T} \right)_V dT + T \left( \frac{\partial S}{\partial V} \right)_T dV = \left( \frac{\partial U}{\partial T} \right)_V dT + \left[ \left( \frac{\partial U}{\partial V} \right)_T + P \right] dV$$

so

$$\left( \frac{\partial S}{\partial T} \right)_V = \frac{1}{T} \left( \frac{\partial U}{\partial T} \right)_V = \frac{C_V}{T}; \quad \left( \frac{\partial S}{\partial V} \right)_T = \frac{1}{T} \left[ \left( \frac{\partial U}{\partial V} \right)_T + P \right] \quad (6-8)$$

Now apply the highly useful equation (3-7), on these partials of  $S$ ,

$$\frac{1}{T} \left[ \frac{\partial}{\partial V} \left( \frac{\partial U}{\partial T} \right)_V \right]_T = \frac{1}{T} \left[ \frac{\partial}{\partial T} \left( \frac{\partial U}{\partial V} \right)_T \right]_V + \frac{1}{T} \left( \frac{\partial P}{\partial T} \right)_V - \frac{1}{T^2} \left[ \left( \frac{\partial U}{\partial V} \right)_T + P \right]$$

or

$$T \left( \frac{\partial P}{\partial T} \right)_V = \left( \frac{\partial U}{\partial V} \right)_T + P = T \left( \frac{\partial S}{\partial V} \right)_T \quad (6-9)$$

We also see that

$$(C_V / \partial V)_T = T \left[ \frac{\partial}{\partial T} \left( \frac{\partial S}{\partial V} \right)_T \right]_V = T (\partial^2 P / \partial T^2)_V \quad (6-10)$$

for any substance having only  $V$  and  $T$  as variables.

From these relationships we can compute the change of  $T$  with  $V$  at constant  $U$ , by using Eq. (4-4) as well as Eq. (3-6) again:

$$\left( \frac{\partial T}{\partial V} \right)_U = - \frac{(\partial U / \partial V)_T}{(\partial U / \partial T)_V} = \frac{1}{C_V} \left[ P - T \left( \frac{\partial P}{\partial T} \right)_V \right] \quad (6-11)$$

This can be integrated if we know the empirical formulas for  $C_V$  and for  $P$  in terms of  $T$  and  $V$ . For example, for a perfect gas,  $P = nRT/V$ , so  $P - T(\partial P / \partial T)_V = 0$  and therefore  $(\partial T / \partial V)_U = 0$ . A perfect gas comes to equilibrium, after a free expansion in an insulated container, with no change in temperature. Joule first proposed and carried out measurements on such processes, to see how closely actual gases behave like perfect gases. For monatomic gases the state function  $n(\partial T / \partial V)_U$  (which is called the Joule coefficient) is less than 0.001°K. moles per  $m^3$ .

For a gas satisfying the Van der Waals equation (3-2), having a heat capacity  $C_V$  that is independent of  $V$ ,

$$P - T \left( \frac{\partial P}{\partial T} \right)_V = - \frac{an^2}{V^2}$$

so

$$\left( \frac{\partial T}{\partial V} \right)_U = \frac{-an^2}{V^2 C_V(T)}$$

which is small because  $a$  is small for most gases. This means that we can consider  $C_V$  constant over the limited range of temperature involved, in which case the total temperature change caused by free expansion is

$$T_1 - T_0 \simeq - \frac{an^2}{C_V} \left( \frac{1}{V_0} - \frac{1}{V_1} \right) \quad (6-12)$$

which gives the small change in  $T$  during the expansion at constant  $U$  for a Van der Waals gas. Since  $V_1 > V_0$  the temperature drops during the process, although the drop is small. During the expansion a small amount of the molecular kinetic energy must be lost in doing work against the small attractive forces between the molecules.

### Entropy of a Gas

To compute the change of  $S$  with  $V$  at constant  $U$  for any substance we need only note that the equation  $T dS = dU + P dV$  gives us directly

$$(\partial S / \partial V)_U = P/T \quad (6-13)$$

which, for a perfect gas, results in  $S_1 - S_0 = nR \ln(V_1/V_0)$  [which could have been obtained from Eq. (6-5)]. For a Van der Waals gas,

$$\left( \frac{\partial S}{\partial V} \right)_U = \frac{nR}{V - nb} - \frac{an^2}{V^2 T(V)}$$

which can be integrated after we have substituted for  $T$  from Eq. (6-11) (change  $T_1, V_1$  into  $T, V$  and solve for  $T$  as a function of  $V, V_0, T_0$ , then substitute this for the  $T$  in the term  $an^2/V^2 T$  and integrate). However, as we found from Eq. (6-12),  $T$  changes but little during the free expansion, so little error is made by setting  $T = T_0$  in the small term ( $an^2/V^2 T$ ). Thus a good approximation to the change in entropy of a Van der Waals gas during change from  $V_0$  to  $V_1$  at constant  $U$  is

$$S_1 - S_0 \simeq nR \ln \left( \frac{V_1 - nb}{V_0 - nb} \right) - \frac{an^2}{T_0} \left( \frac{1}{V_0} - \frac{1}{V_1} \right) \quad (6-14)$$

which should be compared with the entropy change for the perfect gas. The entropy increases in both cases, as it must for a spontaneous process of this sort. Since the system is insulated during the process, the change represents an increase in entropy of the universe.

If we were to go from state 0 to state 1 by a reversible process, the increase in entropy of the gas would be the same as the value we have just calculated. But, to offset this, some other system would lose an equal amount, so the entropy change of the universe would be zero. For example, a reversible change from  $V_0, T_0$  to  $V_1, T_0$  for a perfect gas would be an isothermal expansion, replacing the diaphragm by a piston and moving the piston slowly from  $V_0$  to  $V_1$ . The gas would do work  $W = nRT_0 \ln(V_1/V_0)$  and would gain entropy [see Eq. (6-9)]

$$\Delta S = \int_{V_0}^{V_1} (\partial S / \partial V)_T dV = nR \ln(V_1/V_0)$$

which is the same as that gained during the irreversible process. For the isothermal process, however, we have a heat reservoir at  $T_0$  attached, which loses the amount of entropy that the gas gains, so the change in entropy in the universe is zero in this case.

### Entropy of Mixing

As a final example, illustrating the connection between entropy and disorder, we shall demonstrate that mixing two gases increases their combined entropy. As shown in Fig. 6-3, we start with  $\alpha n$  moles of a gas of type 1 (such as helium) in a volume  $\alpha V$  at temperature  $T_0$  and

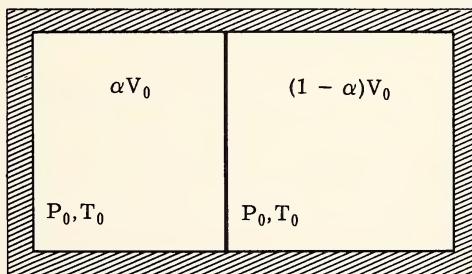


FIG. 6-3. Arrangement to illustrate spontaneous mixing of two different gases.

pressure  $P_0 = \alpha nRT_0/\alpha V$  on one side of a diaphragm and  $(1 - \alpha)n$  moles of gas of type 2 (such as nitrogen) in volume  $(1 - \alpha)V$  on the other side of the diaphragm, in equilibrium at the same temperature and pressure. We now destroy the diaphragm and let the gases spontaneously mix. According to our earlier statements the entropy should increase, since the mixing is an irreversible process of an isolated system. The total internal energy  $U$  of the combined system finally has the same value as the sum of the  $U$ 's of the two parts initially. What has happened is that the type 1 gas has expanded its volume from  $\alpha V$  to  $V$  and type 2 gas has expanded from  $(1 - \alpha)V$  to  $V$ .

To compute the change of entropy, we use  $(\partial S / \partial V)_U$ , which Eq. (6-12) shows is  $P/T$ , which is  $nR/V$  for a perfect gas. Thus the increase of entropy of gas 1 is  $\alpha nR$  times the logarithm of the ratio between final and initial volumes occupied by gas 1, with a similar expression for gas 2. The total entropy increase, called the *entropy of mixing* of the two gases, is

$$\Delta S = nR \left[ \alpha \ln \frac{1}{\alpha} + (1 - \alpha) \ln \left( \frac{1}{1 - \alpha} \right) \right] \quad (6-15)$$

which is positive for  $0 < \alpha < 1$ . It is largest for  $\alpha = 1/2$ , when equal mole quantities of the two gases are mixed. We note that for perfect gases,  $P$  and  $T$  of the final state are the same as those of the initial state; for nonperfect gases  $P$  and  $T$  change somewhat during the mixing (why?).

Entropy increase is to be expected when two *different* gases are mixed. But what if the two gases are the same? Does the removal of a diaphragm separating two parts of a volume  $V$ , filled with one sort of gas, change its entropy or not? When the diaphragm is in place a molecule on one side of it is more restricted in its travel than when the diaphragm is removed; but this difference is unnoticeable macroscopically. Does reinsertion of the diaphragm reduce the entropy again? We must postpone the resolution of this paradox (called Gibbs' paradox) until we treat statistical mechanics (Chapter 22).

# 7

# Simple Thermodynamic Systems

We have already worked out the thermal properties of a perfect gas. According to Eq. (6-9) its total energy is independent of volume for, since  $P = nRT/V$ ,

$$(\partial U / \partial V)_T = T(\partial P / \partial T)_V - P = 0$$

We note in passing that the  $T$  in Eq. (6-9) is the thermodynamic temperature, so that the  $T$  in the perfect gas law  $PV = nRT$  is likewise the thermodynamic temperature. Thus our definition of a perfect gas must include the statement that the  $T$  in the equation of state (3-1) is the thermodynamic temperature.

### The Joule-Thomson Experiment

To see how nearly actual gases come to perfect gases in behavior we can use the free expansion illustrated in Fig. 6-2. The measurement of temperature change in free expansion is called Joule's experiment and the partial derivative which is thus measured,  $n(\partial T / \partial V)_U$ , is called the *Joule coefficient* of the gas. We note from Eqs. (6-10) and (6-11) that the Joule coefficient for a perfect gas is zero and that for a gas obeying Van der Waal's equation it is  $-an^2/V^2C_V$ , a small quantity for most gases.

As with the transformation of heat into work, we can devise a continuous process corresponding to free expansion, as illustrated in Fig. 7-1. Gas is forced through a nozzle  $N$  by moving pistons  $A$  and  $B$  to maintain a constant pressure difference across the nozzle (or we can use pumps working at the proper rates). The gas on the high-pressure side is at pressure  $P_0$  and temperature  $T_0$ ; after going through the nozzle it settles down to a pressure  $P_1$  and temperature  $T_1$ . Suppose we follow  $n$  moles of the gas as it goes through the nozzle, starting from a state of equilibrium 0 at  $P_0, T_0$  and ending at the state of equilibrium 1 at  $P_1, T_1$ . We cannot follow the change in detail, for the process is irreversible, but we can devise a reversible path between

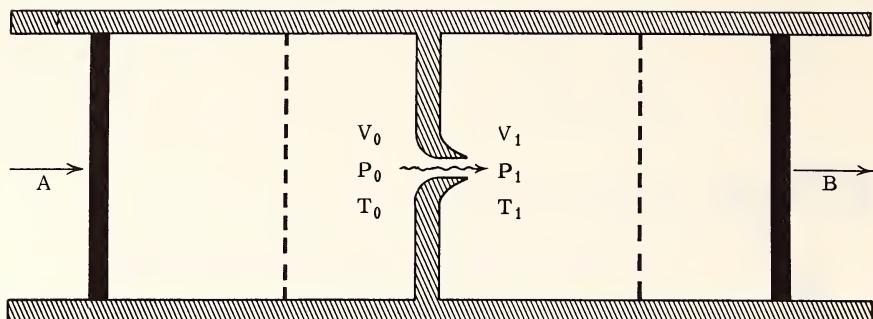


FIG. 7-1. The Joule-Thompson experiment.

0 and 1 [or, rather, we can let Eq. (6-7) find a reversible path for us] which will allow us to compute the difference between states 0 and 1. In particular, we can compute the temperature difference  $T_1 - T_0$  in order to compare it with the measured difference.

This process differs from free expansion because the energy  $U$  does not stay constant; work  $P_0 V_0$  is done by piston A in pushing the gas through the nozzle, and work  $P_1 V_1$  is done on piston B by the time the  $n$  moles have all gotten through. Thus the net difference in internal energy of the  $n$  moles between state 0 and state 1 is  $U_1 - U_0 = P_0 V_0 - P_1 V_1$ . Instead of  $U$  remaining constant during the process, the quantity

$$H = U + PV \quad (7-1)$$

is the same in states 1 and 0. This quantity, called the *enthalpy* of the gas, is an extensive state variable.

The experiment of measuring the difference  $T_1 - T_0$ , when all parts of the system shown in Fig. 7-1 are thermally insulated, is called the Joule-Thompson experiment, and the relevant change in temperature with pressure when  $H$  is kept constant,  $(\partial T / \partial P)_H$ , is the *Joule-Thompson coefficient* of the gas. Experimentally, we find that for actual gases, the temperature increases slightly at high temperatures; at low temperatures the temperature  $T_1$  is a little less than  $T_0$ . The temperature at which  $(\partial T / \partial P)_H = 0$  is called the Joule-Thompson *inversion point*. Continuous processes, in which a gas, at a temperature below its inversion point, is run through a nozzle to lower its temperature, are used commercially to attain low temperature. We use work  $P_0 V_0 - P_1 V_1$  to cool the  $n$  moles of gas, so Clausius' principle is not contradicted.

To compute the Joule-Thompson coefficient we manipulate the equation for  $H$ , or rather its differential,

$$dH = dU + P dV + V dP = T dS + V dP \quad (7-2)$$

as we did in Eq. (6-7) to obtain Eq. (6-10). First we note that the change in heat  $T dS$  in a system at constant pressure equals  $dH$ , so that  $C_p$ , the heat capacity of the system at constant pressure, is  $(\partial H / \partial T)_p$ , in contrast to Eq. (4-4). Just as internal energy  $U$  can be called the heat content of a system at constant volume, so enthalpy  $H$  can be called its heat content at constant pressure. In passing, we can combine Eqs. (4-6) and (6-9) to express the difference between  $C_p$  and  $C_v$  for any system in terms of partials obtainable from its equation of state,

$$C_p - C_v = T(\partial V / \partial T)_p (\partial P / \partial T)_v \quad (7-3)$$

Next we manipulate Eq. (7-2) as we did Eq. (6-7), to obtain

$$\left(\frac{\partial S}{\partial T}\right)_p = \frac{1}{T} \left(\frac{\partial H}{\partial T}\right)_p = \frac{C_p}{T}; \quad \left(\frac{\partial S}{\partial P}\right)_T = \frac{1}{T} \left[ \left(\frac{\partial H}{\partial P}\right)_T - V \right] = - \left(\frac{\partial V}{\partial T}\right)_p \quad (7-4)$$

and

$$(\partial C_p / \partial P)_T = -T(\partial^2 V / \partial T^2)_p$$

From this we can compute the Joule-Thomson coefficient and the change in entropy during the process,

$$\left(\frac{\partial T}{\partial P}\right)_H = - \frac{(\partial H / \partial P)_T}{(\partial H / \partial T)_p} = - \frac{1}{C_p} \left[ V - T \left(\frac{\partial V}{\partial T}\right)_p \right] \quad \text{and} \quad \left(\frac{\partial S}{\partial P}\right)_H = - \frac{V}{T} \quad (7-5)$$

For a perfect gas  $V = T(\partial V / \partial T)_p$ , so that  $(\partial H / \partial P)_T = 0$  and the Joule-Thomson coefficient  $(\partial T / \partial P)_H$  is also zero; no change in temperature is produced by pushing it through a nozzle. The change in entropy of a perfect gas during the process is the integral of  $-V/T = -nR/P$  with respect to  $P$ ,

$$\Delta S = nR \ln(P_0/P_1) \quad (7-6)$$

Since  $P_0 > P_1$  this represents an increase in entropy, as it must.

For a gas obeying Van der Waal's equation we can find  $(\partial V / \partial T)_p$  by differentiating the equation of state and manipulating,

$$(V - nb) dP + \left[ \frac{nRT}{V - nb} - \frac{2an^2}{V^3} (V - nb) \right] dV = nR dT$$

$$T \left(\frac{\partial V}{\partial T}\right)_p = (V - nb) \left[ 1 - \frac{2an}{V^3} \frac{(V - nb)^2}{RT} \right]^{-1} \simeq (V - nb) \left( 1 + \frac{2an}{RTV} \right)$$

so that

$$\left(\frac{\partial T}{\partial P}\right)_H \simeq \frac{+1}{C_p} \left( \frac{2an}{RT} - nb \right)$$

since  $a$  and  $b$  are small quantities. For this same reason  $T$  and thus  $C_p$  do not change much during the process, and we can write

$$T_1 - T_0 \simeq \frac{nb}{C_p} \left( 1 - \frac{2a}{RT_0 b} \right) (P_0 - P_1)$$

Since  $P_0 > P_1$ , this predicts an increase in temperature during the Joule-Thomson process if  $T_0 > 2a/Rb$ , a decrease if  $T_0 < 2a/Rb$ , the inversion temperature being approximately equal to  $2a/Rb$ .

The corresponding increase in entropy during the process can be computed, using the same approximations as before,

$$\Delta S \simeq nR \ln\left(\frac{P_0}{P_1}\right) - \frac{nb}{T_0} (P_0 - P_1) \left( 1 - \frac{a}{RT_0 b} \right) \quad (7-7)$$

which is to be compared with the result of Eq. (7-6) for a perfect gas.

### Black-Body Radiation

Let us now turn to a quite different sort of system, that called black-body radiation, electromagnetic radiation in equilibrium with the walls of an enclosure kept at temperature  $T$ . This is the radiation one would find inside a furnace with constant-temperature walls. It consists of radiation of all frequencies and going in all directions, some of it continually being absorbed by the furnace walls but an equal amount continually being generated by the vibrations of the atoms in the walls. This is a special kind of system, with special properties.

In the first place the energy density  $e$  of the radiation, the mean value of  $(1/2) \varepsilon \cdot \mathcal{D} + (1/2) \mathcal{K} \cdot \mathcal{G}$ , depends on the temperature but is independent of the volume of the enclosure. If the volume inside the furnace is enlarged, more radiation is generated by the walls, so that the energy density  $e$  remains the same. Therefore the total electromagnetic energy within the enclosure,  $V_e(T)$ , is proportional to the volume. This is in contrast to a perfect gas, for which  $U$  at a given temperature is constant, independent of volume. If the enclosure is increased in volume the density of atoms diminishes and so does the energy density; for radiation, if the volume is increased more radiation is produced to keep the density constant. We can, of course, consider the radiation to be a gas of photons, each with its own energy, but the contrast with atoms remains. Extra photons can be created to fill up any added space; atoms are harder to create.

In the second place the radiation pressure, the force exerted per unit area on the container walls, is proportional to the energy density. In this respect it is similar to a perfect gas of point atoms, but the proportionality constant is different. For the gas [see Eq. (2-4)]  $P = (2/3)[N < K.E. >_{\text{tran}}/V] = (2/3)e$ , where  $e$  is the energy contained per unit volume; for radiation it is  $(1/3)e$ . The difference comes from the fact that the kinetic energy of an atom,  $(1/2)mv^2$ , is *one-half* its momentum times its velocity. For a photon, if its energy is  $\hbar\omega$ , its momentum is  $(\hbar\omega/c)$  where  $c$  is its velocity,  $(\omega/2\pi)$  its frequency, and  $\hbar = 2\pi\hbar$  is Planck's constant; therefore the energy of a photon is the product of its velocity and its momentum, not half this product. Since pressure is proportional to momentum,  $e = 3P$  for the photon,  $= (3/2)P$  for the atom gas.

To compute  $U$ ,  $S$ , and  $P$  as functions of  $T$  and  $V$  we start with the basic equation again,  $dU = T dS - P dV$ , inserting the appropriate expressions,  $U = Ve(T)$  and  $P = (1/3)e(T)$ ,

$$\begin{aligned} T dS &= T \left( \frac{\partial S}{\partial T} \right)_V dT + T \left( \frac{\partial S}{\partial V} \right)_T dV = dU + P dV \\ &= V \left( \frac{de}{dT} \right) dT + e dV + \frac{1}{3} e dV \end{aligned}$$

or

$$\left( \frac{\partial S}{\partial T} \right)_V = \frac{V}{T} \left( \frac{de}{dT} \right); \quad \left( \frac{\partial S}{\partial V} \right)_T = \frac{4}{3} \frac{e}{T}$$

Applying Eq. (3-10) we obtain a differential equation for  $e(T)$ ,

$$\frac{1}{T} \left( \frac{de}{dT} \right) = \frac{4}{3} \frac{1}{T} \left( \frac{de}{dT} \right) - \frac{4}{3} \frac{e}{T^2} \quad \text{or} \quad \frac{de}{dT} = 4 \frac{e}{T}$$

which has a solution  $e(T) = aT^4$ , so

$$U = aVT^4; \quad P = \frac{1}{3}aT^4; \quad S = \frac{4}{3}aVT^3 \quad (7-8)$$

The equation for the energy density of black-body radiation is called *Stefan's law* and the constant  $a$  is Stefan's constant. In statistical mechanics [see Eq. (25-9)] we shall evaluate it in terms of atomic constants.

We see that the energy of black-body radiation goes up very rapidly with increase in temperature. Room temperature ( $70^\circ \text{F}$ ) is about  $300^\circ \text{K}$ . At the temperature of boiling water ( $373^\circ \text{K}$ ) the energy density of radiation is already  $2 \frac{1}{2}$  times greater; at dull red heat ( $920^\circ \text{K}$ ) it

is 100 times that at room temperature. At the temperatures encountered on earth the pressure of radiation is minute compared to usual gas pressures. At temperatures of the center of the sun ( $10^7$  °K) the radiation pressure supports more than half the mass above it.

Reference to Eq. (6-3),  $U = ST - PV + \mu n$  shows that the chemical potential of black-body radiation is

$$\mu = (1/n)(U - ST + PV) = (1/n) \left( aVT^4 - \frac{4}{3} aVT^4 + \frac{1}{3} aVT^4 \right) = 0$$

which is related to the freedom with which the number of moles of photons adjusts its value to keep the energy density constant [see Eq. (25-2)]. In this matter, also, photons are a special kind of gas.

### Paramagnetic Gas

Finally let us work out the behavior of a paramagnetic, perfect gas. Here the two equations of state are  $P = nRT/V$  and  $\mathcal{C} = T\mathfrak{M}/nD$ . The heat capacity at constant  $V$  and  $\mathfrak{M}$  is a constant.  $C_{V\mathfrak{M}} = (3/2)nR$  for a monatomic gas. The basic equation is

$$dU = T dS - P dV + \mathcal{C} d\mathfrak{M}$$

Three independent variables must be used; we first use  $T$ ,  $V$ , and  $\mathfrak{M}$ . By methods that should be familiar by now we find

$$\begin{aligned} T \left( \frac{\partial S}{\partial T} \right)_{V\mathfrak{M}} &= \frac{3}{2} nR = \left( \frac{\partial U}{\partial T} \right)_{V\mathfrak{M}} \\ T \left( \frac{\partial S}{\partial V} \right)_{T\mathfrak{M}} &= P + \left( \frac{\partial U}{\partial V} \right)_{T\mathfrak{M}} = T \left( \frac{\partial P}{\partial T} \right)_V \quad (7-9) \\ T \left( \frac{\partial S}{\partial \mathfrak{M}} \right)_{TV} &= \left( \frac{\partial U}{\partial \mathfrak{M}} \right)_{TV} - \mathcal{C} = -T \left( \frac{\partial \mathcal{C}}{\partial T} \right)_{\mathfrak{M}} \end{aligned}$$

Other manipulations result in

$$(\partial U / \partial V)_{T\mathfrak{M}} = (\partial U / \partial \mathfrak{M})_{TV} = 0$$

$$C_{P\mathcal{C}} = C_{V\mathfrak{M}} + P \left( \frac{\partial V}{\partial T} \right)_P - \mathcal{C} \left( \frac{\partial \mathfrak{M}}{\partial T} \right)_{\mathcal{C}} = \frac{5}{2} nR + (\mathfrak{M}^2 / nD)$$

Integrating the partials for  $U$  and  $S$ , we obtain

$$U = \frac{3}{2} nRT; \quad S = nR \ln \left[ \frac{V}{V_0} \left( \frac{T}{T_0} \right)^{3/2} \right] - \frac{\mathfrak{M}^2}{2nD} + S_0 \quad (7-10)$$

The "natural variables" in terms of which to express  $U$  are  $S$ ,  $V$ ,  $\mathfrak{M}$  rather than  $T$ ,  $V$ ,  $\mathfrak{M}$ . To do this we express  $T$  in terms of  $S$ ,  $V$ ,  $\mathfrak{M}$  and obtain

$$U = nRT_0 \left( \frac{V_0}{V} \right)^{2/3} \exp \left( \frac{2}{3} \frac{S - S_0}{nR} + \frac{\mathfrak{M}^2}{3n^2RD} \right) \quad (7-11)$$

[see Eq. (6-5) et seq.].

The extensive variables  $S$ ,  $V$ ,  $\mathfrak{M}$  are less easy to measure than are the intensive variables

$$P = -(\partial U / \partial V)_{S\mathfrak{M}}; \quad T = (\partial U / \partial S)_{V\mathfrak{M}}; \quad H = (\partial U / \partial \mathfrak{M})_{SV}$$

which might be called the "experimental variables." Expressing the basic equation in terms of these (using  $P dV = -(nRT/P)dP + nR dT$ , for example) we have

$$T dS = \left( \frac{5}{2} nR + nD \frac{\mathfrak{M}^2}{T^2} \right) dT - \frac{nRT}{P} dP - \frac{nD\mathfrak{M}}{T} d\mathfrak{M} \quad (7-12)$$

For isothermal operation,  $dT$  is zero. The heat contributed to the gas by the reservoir at temperature  $T$ , when the gas pressure is changed from  $P_0$  to  $P_1$  and the magnetic intensity is changed from  $\mathfrak{M}_0$  to  $\mathfrak{M}_1$ , is

$$\Delta Q_{01} = nRT \ln(P_0/P_1) + (nD/2T)(\mathfrak{M}_0^2 - \mathfrak{M}_1^2) \quad (7-13)$$

Increase in pressure squeezes out heat ( $\Delta Q < 0$ ) as does an increase in magnetization. At low temperatures a change in magnetic field produces more heat than does a change of pressure.

The behavior of the system during an adiabatic process can be computed by setting  $T dS = 0$ . The integrating factor is  $1/T$  and the integral of

$$\left( \frac{5}{2} \frac{R}{T} + D \frac{\mathfrak{M}^2}{T^3} \right) dT - \frac{nD\mathfrak{M}}{T^2} d\mathfrak{M} - \frac{nR}{P} dP = 0$$

is

$$\left( \frac{P_0}{P_1} \right) \left( \frac{T_1}{T_0} \right)^{5/2} \exp \left[ \frac{D}{2R} \left( \frac{\mathfrak{M}_0^2}{T_0^2} - \frac{\mathfrak{M}_1^2}{T_1^2} \right) \right] = 1 \quad (7-14)$$

If the magnetic field is kept constant ( $\mathfrak{M}_1 = \mathfrak{M}_0$ ), the gas undergoes ordinary adiabatic compression and  $T_1$  is proportional to the two-fifths power of the pressure  $P_1$ . Or, if the pressure is kept constant, the temperature is related to the magnetic field by the formula

$$\mathcal{K}_1^2 = T_1^2 \left[ \left( \mathcal{K}_0^2 / T_0^2 \right) + \left( 5R/D \right) \ln \left( T_1 / T_0 \right) \right] \quad (7-15)$$

At low-enough initial temperatures or high-enough initial fields, so that  $(\mathcal{K}_0/T_0)^2 \gg 5R/D$ , the final temperature  $T_1$  is approximately proportional to the final magnetic intensity  $\mathcal{K}_1$ ; an adiabatic reduction of  $\mathcal{K}$  proportionally lowers the temperature.

Finally, we can have a process which is *both* adiabatic and isothermal if we adjust the pressure continually so that, as the magnetic field is changed adiabatically, the temperature is kept at a constant value  $T_0$  (the volume then changes inversely proportional to  $P$ ). The relation between  $\mathcal{K}_1$  and  $P_1$  that will keep  $T$  constant is

$$P_1 = P_0 \exp \left[ \frac{D}{2RT_0^2} (\mathcal{K}_0^2 - \mathcal{K}_1^2) \right] \quad (7-16)$$

As  $\mathcal{K}_1$  decreases,  $P_1$  will have to be increased exponentially to keep  $T$  constant. In this process, mechanical work is used to demagnetize the material.

Actual magnetic materials are not perfect gases, nor does their magnetic equation of state have exactly the form of the simple Curie law. In these cases one tries to find equations of state that do fit the data and then tries to integrate Eq. (7-12). Failing this, numerical integration must be used to predict the thermal properties of the material.

# 8 The Thermodynamic Potentials

In connection with Eq. (6-1) we pointed out that  $U$ , the internal energy, when expressed in terms of the extensive variables  $S$ ,  $V$ ,  $L$ ,  $\mathfrak{M}$ ,  $n_i$ , etc., behaves like a potential energy, in that the thermodynamic "forces," the intensive variables, are expressible as gradients of  $U$ ,

$$T = \left( \frac{\partial U}{\partial S} \right)_{vn\dots}; \quad P = - \left( \frac{\partial U}{\partial V} \right)_{Sn\dots}; \quad \mu = \left( \frac{\partial U}{\partial n} \right)_{SV\dots}; \quad \text{etc.} \quad (8-1)$$

as was indicated in Eq. (6-2).

## The Internal Energy

Moreover if we include irreversible as well as reversible processes, Eqs. (4-1), (6-1), and (6-5) show that

$$\begin{aligned} dU &= \bar{d}Q - P dV + \mu dn + J dL + \mathfrak{C} d\mathfrak{M} + \dots \\ &\leq T dS - P dV + \mu dn + J dL + \mathfrak{C} d\mathfrak{M} + \dots \end{aligned} \quad (8-2)$$

where the equality holds for reversible processes, the inequality for irreversible ones. Consequently if *all* the extensive variables are held constant ( $dS = dV = dL = \dots = 0$ ) while the system is allowed to come spontaneously to equilibrium, then every change in  $U$  will have to be a decrease in value; it will only stop changing spontaneously when no further decrease is possible; and thus at equilibrium  $U$  is minimal. In other words, when the extensive variables are fixed, the equilibrium state is the state of minimal  $U$ .

It is usually possible (although not always easy) to hold the mechanical extensive variables,  $V$ ,  $L$ ,  $n$ ,  $\mathfrak{M}$ , etc., constant, but it is more difficult to hold entropy constant during a spontaneous process. However if we allow the process to take place adiabatically, with  $\bar{d}Q = 0$

(and hold  $dV = dL = dn = \dots = 0$ ), then  $U$  does not change during the process and  $T dS \geq 0$ , so that  $S$  will reach a *maximum* value at equilibrium. If we had tried to keep  $S$  constant during the spontaneous process we would have had to withdraw heat during the process, enough to cancel out the gain in  $S$  during the process; this would have decreased  $U$ , until at equilibrium  $U$  would be minimal, less than the original  $U$  by the amount of heat that had to be withdrawn to keep  $S$  constant.

When all the mechanical extensive variables are held constant, an addition of heat to the system produces a corresponding increase in  $U$ . We thus can call  $U$  the heat content of the system at constant  $V$ ,  $L$ ,  $n$ ,  $\dots$ , and can write the heat capacity

$$C_{vn\dots} = (\partial U / \partial T)_{vn\dots} \quad (8-3)$$

the subscripts indicating that all the extensive mechanical variables, which apply to the system in question, are constant. Moreover we can apply Eq. (3-10) to the various intensive variables of Eq. (8-1), obtaining a very useful set of relationships,

$$\begin{aligned} \left(\frac{\partial T}{\partial V}\right)_{Sn\dots} &= - \left(\frac{\partial P}{\partial S}\right)_{Vn\dots}; & \left(\frac{\partial T}{\partial n}\right)_{SV\dots} &= \left(\frac{\partial \mu}{\partial S}\right)_{Vn\dots} \\ \left(\frac{\partial T}{\partial \mu}\right)_{SV\dots} &= \left(\frac{\partial \mathcal{H}}{\partial S}\right)_{V\mathcal{M}\dots}; & - \left(\frac{\partial P}{\partial n}\right)_{SV\dots} &= \left(\frac{\partial \mu}{\partial V}\right)_{Sn\dots} \\ - \left(\frac{\partial P}{\partial \mathcal{M}}\right)_{SV\dots} &= \left(\frac{\partial \mathcal{H}}{\partial V}\right)_{S\mathcal{M}\dots}; & \text{etc.} & \end{aligned} \quad (8-4)$$

which are called *Maxwell's relations*. The relations, of course, hold for reversible processes. They are used to compute the differences in value of the various thermodynamic variables between two equilibrium states.

### Enthalpy

Although the extensive variables are the "natural" ones in which to express  $U$ , they are often not the most useful ones to use as independent variables. We do not usually measure entropy directly, we measure  $T$ ; and often it is easier to keep pressure constant during a process than it is to keep  $V$  constant. We should look for functions that have some or all of the intensive variables as the "natural" ones. Formally, this can be done as follows. Suppose we wish to change from  $V$  to  $P$  for the independent variable. We add the product  $PV$  to  $U$ , to generate the function  $H = U + PV$ , called the enthalpy [see Eq. (7-1)]. The differential is

$$dH = P dV + V dP + dU$$

$$\leq T dS + V dP + \mu dn + \mathcal{H} d\mathcal{M} + \dots \quad (8-5)$$

where again the equality holds for reversible processes, the inequality for irreversible ones. If a system is held at constant  $S, P, n, \dots$ , the enthalpy will be minimal at equilibrium (note that  $P$ , instead of  $V$ , is held constant). If  $H$  is expressed as a function of its "natural" coordinates  $S, P, n, \dots$ , then the partials of  $H$  are the quantities

$$\left(\frac{\partial H}{\partial S}\right)_{Pn\dots} = T; \quad \left(\frac{\partial H}{\partial P}\right)_{Sn\dots} = V; \quad \left(\frac{\partial H}{\partial n}\right)_{SP\dots} = \mu; \quad \text{etc.} \quad (8-6)$$

and the corresponding Maxwell's relations are

$$\left(\frac{\partial T}{\partial P}\right)_{Sn\dots} = \left(\frac{\partial V}{\partial S}\right)_{Pn\dots}; \quad \left(\frac{\partial V}{\partial \mathcal{M}}\right)_{SP\dots} = \left(\frac{\partial \mathcal{H}}{\partial P}\right)_{S\mathcal{M}\dots}; \quad \text{etc.} \quad (8-7)$$

Geometrically, the transformation from  $U$  to  $H$  is an example of a *Legendre transformation* involving the pair of variables  $P, V$ . Function  $U(V)$ , for a specific value of  $V$  (i.e., at point  $Q$  in Fig. 8-1) has

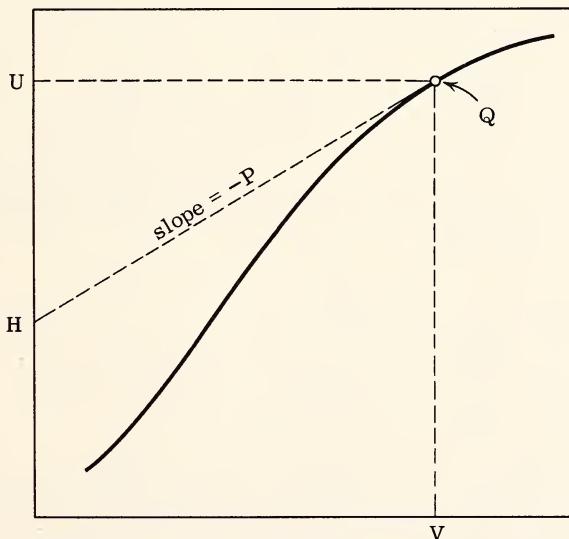


FIG. 8-1. Legendre transformation from  $U$  as a function of  $V$  to  $H$  as a function of  $P$

a slope  $dU/dV = -P(V)$ , which defines a tangent,  $HQ$  of Fig. 8-1, which has an intercept on the  $U$  axis of  $H = U + PV$ . (We are not con-

sidering any other variables except  $P$ ,  $V$ , so we can use ordinary derivatives for the time being.) Solving for  $V$  as a function of  $P$  from the equation  $dU/dV = -P$ , we can then express  $H$  as a function of the slope  $-P$  of the tangent line. Since  $dU = -P dV$  and  $dH = dU + P dV + V dP = V dP$  we see that  $dH(P)/dP = V(P)$ . Thus enthalpy  $H$  is the potential that has  $P$  as a basic variable instead of  $V$ .

Since any addition of heat  $T dS$  to a system when  $P, n, \dots$  are held constant causes a like increase of  $H$ , the enthalpy can be called the heat content of the system at constant pressure.

### The Helmholtz and Gibbs Functions

At least as important as the change from  $V$  to  $P$  is the change from  $S$  to  $T$ . Adiabatic (constant  $S$ ) processes are encountered, of course, but many thermodynamic measurements are carried out at constant temperature. The Legendre transformation appropriate for this produces the *Helmholtz function*  $F = U - TS$ , for which

$$dF \leq -S dT - P dV + J dL + \mu dn + \mathcal{C} d\mathfrak{M} + \dots \quad (8-8)$$

$$\left(\frac{\partial F}{\partial T}\right)_{Vn\dots} = -S; \quad \left(\frac{\partial F}{\partial V}\right)_{Tn\dots} = -P; \quad \left(\frac{\partial F}{\partial \mathfrak{M}}\right)_{TV\dots} = \mathcal{C}; \quad \text{etc.}$$

The related Maxwell relations include

$$\left(\frac{\partial S}{\partial V}\right)_{Tn\dots} = \left(\frac{\partial P}{\partial T}\right)_{Vn\dots}; \quad \left(\frac{\partial S}{\partial \mathfrak{M}}\right)_{TV\dots} = -\left(\frac{\partial \mathcal{C}}{\partial T}\right)_{V\mathfrak{M}\dots}; \quad \text{etc.} \quad (8-9)$$

Since the heat capacity  $C_{V\mathfrak{M}\dots}$  is equal to  $T(\partial S/\partial T)_{V\mathfrak{M}\dots}$  [see Eq. (6-8)], by differentiating these relations again with respect to  $T$  we obtain a set of equations

$$\begin{aligned} \left(\frac{\partial}{\partial V} C_{V\mathfrak{M}\dots}\right)_{T\mathfrak{M}\dots} &= T \left(\frac{\partial^2 P}{\partial T^2}\right)_V \\ \left(\frac{\partial}{\partial \mathfrak{M}} C_{V\mathfrak{M}\dots}\right)_{TV\dots} &= -T \left(\frac{\partial^2 \mathcal{C}}{\partial T^2}\right)_{\mathfrak{M}}; \quad \text{etc.} \end{aligned} \quad (8-10)$$

which indicate that the heat capacity is not completely independent of the equations of state. The dependence of  $C_{V\mathfrak{M}\dots}$  on  $T$  is something we must obtain directly by measurement; its dependence on the other variables,  $V, \mathfrak{M}$ , etc., can be obtained from the equations of state.

By the same arguments as before, the inequality of Eq. (8-8) shows that for a system in which  $T, V, n, \mathfrak{M}$ , etc., are held constant, the Helmholtz function  $F$  is minimal at equilibrium. If  $T$  is held constant, any change in  $F$  can be transformed completely into work, such

as  $-P dV$  or  $\mathcal{M} d\mathcal{M}$ , etc. Thus  $F$  is sometimes called the free energy of the system at constant temperature. The Maxwell relations (8-9) are particularly useful. For example, the second equation shows that, since an increase of magnetic polarization increases the orderliness in orientation of the atomic magnets and hence entropy (a measure of disorder) will decrease as  $\mathcal{M}$  increases, therefore the magnetic intensity  $\mathcal{M}$  required to produce a given magnetization  $\mathcal{M}$  will increase as  $T$  is increased. Similarly we can predict that since rubber tends to change from amorphous to crystal structure (i.e., becomes more ordered) as it is stretched, therefore the tension in a rubber band held at constant length will increase as  $T$  is increased.

Another potential of considerable importance is the *Gibbs function*  $G = U + PV - TS = F + PV$ , having  $T$  and  $P$  for its natural coordinates, instead of  $S$  and  $V$ , and for which

$$dG \leq -S dT + V dP + \mu dn + J dL + \mathcal{M} d\mathcal{M} + \dots \quad (8-11)$$

$$\left(\frac{\partial G}{\partial T}\right)_{Pn\dots} = -S; \quad \left(\frac{\partial G}{\partial P}\right)_{Tn\dots} = V; \quad \text{etc.}$$

The Gibbs function is a minimum at equilibrium for a system held at constant  $T$ ,  $P$ ,  $n$ ,  $\mathcal{M}$ ,  $\dots$ , a property that we will utilize extensively in the next two chapters. The Maxwell relations include

$$\left(\frac{\partial S}{\partial P}\right)_{Tn\dots} = -\left(\frac{\partial V}{\partial T}\right)_{Pn\dots}; \quad \left(\frac{\partial V}{\partial n}\right)_{TP\dots} = \left(\frac{\partial \mu}{\partial P}\right)_{Tn\dots}; \quad \text{etc.} \quad (8-12)$$

from which we can obtain the dependence of the heat capacity  $C_{P\mathcal{M}\dots}$  on the mechanical variables

$$\left(\frac{\partial}{\partial P} C_{P\mathcal{M}\dots}\right)_{T\mathcal{M}\dots} = -T \left(\frac{\partial^2 V}{\partial T^2}\right)_P; \quad \text{etc.} \quad (8-13)$$

This process can be continued for all the mechanical variables, obtaining a new potential each time, and a new set of Maxwell relations. For example there is a *magnetic Gibbs function*  $G_m = U - TS + PV - \mathcal{M} \mathcal{M}$ . From it we can obtain the following:

$$\left(\frac{\partial S}{\partial \mathcal{M}}\right)_{TP\dots} = \left(\frac{\partial \mathcal{M}}{\partial T}\right)_{P\mathcal{M}\dots}; \quad \left(\frac{\partial V}{\partial \mathcal{M}}\right)_{TP\dots} = -\left(\frac{\partial \mathcal{M}}{\partial P}\right)_{T\mathcal{M}\dots}; \quad \text{etc.} \quad (8-14)$$

Finally there is what is called the *grand potential*  $\Omega$ , obtained by a Legendre transformation from  $n$  to  $\mu$  as the independent variable, which is useful in the study of systems with variability of number of

particles, such as some quantum systems exhibit:

$$\Omega = U - TS - \mu n; \quad d\Omega \leq -S dT - P dV - n d\mu + J dL + \dots \quad (8-15)$$

from which the following Maxwell relations come:

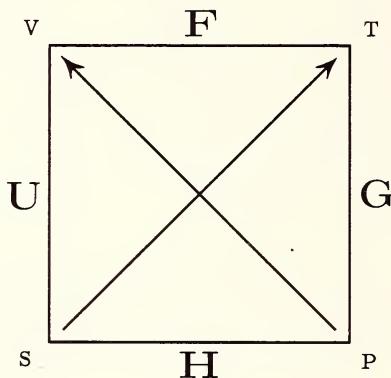
$$\left(\frac{\partial S}{\partial \mu}\right)_{TV\dots} = \left(\frac{\partial n}{\partial T}\right)_{V\mu\dots}; \quad \left(\frac{\partial P}{\partial \mu}\right)_{TV\dots} = \left(\frac{\partial n}{\partial V}\right)_{T\mu\dots}; \text{etc.} \quad (8-16)$$

### Procedures for Calculation

A technique for remembering all these relationships can be worked out for simple systems, such as those definable in terms of  $T$  and  $V$ , with  $n$  constant. Here, when we use Euler's equation (6-3),

$$\begin{aligned} U &= ST - PV + \mu n; & H &= U + PV = ST + \mu n \\ F &= U - ST = -PV + \mu n; & G &= F + PV = \mu n \end{aligned} \quad (8-17)$$

The mnemonic device is shown in Fig. 8-2. It indicates the natural variables for each of the four potentials ( $S$  and  $V$  for  $U$ , etc.) and the arrows indicate the basic partial derivative relations, with their signs,



**FIG. 8-2.** Diagram relating the thermodynamic potentials with their natural variables, their partials, and the partials of these variables (Maxwell relations) for a simple system.

such as  $(\partial H / \partial P)_S = +V$  and  $(\partial F / \partial T)_V = -S$ . It also indicates the nature of the various Maxwell relations, such as  $(\partial S / \partial V)_T = (\partial P / \partial T)_V$

or  $(\partial S/\partial P)_T = -(\partial V/\partial T)_P$ ; the arrows this time connect the numerator of the partial derivative with the subscript on it and the directions of the arrows again indicate the sign.

Using this device we are now in a position to formulate a strategy for expressing any possible rate of change of a thermodynamic variable in terms of the immediately measurable quantities such as heat capacity and an equation of state, or else the empirically determined partials  $(\partial V/\partial T)_P$  and  $(\partial V/\partial P)_T$  from which we could obtain an equation of state. Either  $C_V$  or  $C_P$  can be considered basic ( $C_P$  is the one usually measured),

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V = T \left(\frac{\partial S}{\partial T}\right)_V; \quad C_P = \left(\frac{\partial H}{\partial T}\right)_P = T \left(\frac{\partial S}{\partial T}\right)_P \quad (8-18)$$

The relation between them is given in terms of partials from the equation of state [see Eq. (7-3)],

$$C_P = C_V + T(\partial P/\partial T)_V (\partial V/\partial T)_P \quad (8-19)$$

The various tactics which can be used to express an unfamiliar partial in terms of an immediately measurable one are:

(a) Replacing the partials of the potentials with respect to their adjoining variables in Fig. 8-2 by the related variables, such as  $(\partial F/\partial T)_V = -S$  or  $(\partial U/\partial S)_V = T$ , etc.

(b) Replacing a partial of a potential with respect to a nonadjacent variable, obtainable from its basic equation, such as  $dF = -S dT - P dV$ , from which we can get  $(\partial F/\partial V)_S = -S(\partial T/\partial V)_S - P$ , and  $(\partial F/\partial S)_P = -S(\partial T/\partial S)_P - P(\partial V/\partial S)_P$ , etc.

(c) Using one or more of the Maxwell relations, obtainable from Fig. 8-2.

(d) Using the basic properties of partial derivatives, as displayed in Eqs. (3-9) and (3-10).

In terms of these tactics, the appropriate strategies are:

1. If a potential is an independent variable in the given partial, make it the dependent variable by using (d) (this process is called *bringing the potential into the numerator*) and then use (a) or (b) to eliminate the potential. Examples:

$$\left(\frac{\partial T}{\partial U}\right)_V = \frac{1}{(\partial U/\partial T)_V} = \frac{1}{C_V}$$

$$-\left(\frac{\partial T}{\partial V}\right)_U = \frac{(\partial U/\partial V)_T}{(\partial U/\partial T)_V} = \frac{T}{C_V} \left(\frac{\partial S}{\partial V}\right)_T - \frac{P}{C_V}$$

2. Next, if the entropy is an independent variable in the given par-

tial or in the result of step 1, bring  $S$  into the numerator and eliminate it by using (c) or Eq. (8-18). Examples:

$$\left(\frac{\partial P}{\partial S}\right)_V = \frac{(\partial P/\partial T)_V}{(\partial S/\partial T)_V} = \frac{T}{C_V} \left(\frac{\partial P}{\partial T}\right)_V$$

$$-\left(\frac{\partial T}{\partial V}\right)_S = \frac{(\partial S/\partial V)_T}{(\partial S/\partial T)_V} = \frac{T}{C_V} \left(\frac{\partial P}{\partial T}\right)_V$$

3. If the measured equation-of-state partials have  $V$  in the numerator, bring  $V$  into the numerator in the result of steps 1 and 2 by using (d). If the equation of state is in the form  $P = f(V, T)$ , bring  $P$  into the numerator.

The result of applying these successive steps will be an expression for the partial of interest in terms of measured (or measurable) quantities. It is obvious that if other variables, such as  $m$  and  $c$  instead of  $V$  and  $P$ , are involved, a square similar to Fig. 8-2 can be constructed, which would represent the relationships of interest, and procedures 1 to 4 can be carried through as before.

### Examples and Useful Formulas

As one example, we shall anticipate a problem discussed in Chapter 9. There we shall find it enlightening to plot the Gibbs function  $G$  as a function of  $T$  for constant  $P$ . We can obtain such plots by integrating  $(\partial G/\partial T)_P = -S$  or, since  $S$  is not a directly measured quantity, by the double integration of  $-(\partial S/\partial T)_P = -C_P/T$  with respect to  $T$ . In the case of a gas, where  $C_V$  is a function of  $T$  alone, we must then express  $C_P$  as a function of  $T$  and  $P$ , so it may be integrated easily. Thus several applications of the procedures outlined above result in

$$\left(\frac{\partial^2 G}{\partial T^2}\right)_P = -\frac{C_P}{T} = -\frac{C_V(T)}{T} - \left(\frac{\partial V}{\partial T}\right)_P \left(\frac{\partial P}{\partial T}\right)_V = -\frac{C_V}{T} + \frac{[(\partial V/\partial T)_P]^2}{(\partial V/\partial P)_T}$$

so

$$G = G_0(P) - (T - T_0)S_0(P) - \int_{T_0}^T dT \int_{T_0}^T \left\{ \frac{C_V}{T} - \frac{[(\partial V/\partial T)_P]^2}{(\partial V/\partial P)_T} \right\} dT \quad (8-20)$$

Since  $S_0(P)$ , the entropy at temperature  $T_0$  at the constant  $P$  is a positive quantity; the plot of  $G$  against  $T$  for constant  $P$  has a negative slope. Furthermore, since  $C_V$  is positive and  $(\partial V/\partial P)_T$  is negative for all  $T$  and  $P$ , the slope of  $G$  becomes more negative as  $T$  increases; the plot curves downward (see Figs. 9-2 and 9-6).

Finally, and in part to collect some formulas for useful reference, let us work out again the entropy and then the thermodynamic potentials for a perfect gas of point atoms, from its equation of state  $PV = nRT$  and its heat capacity  $C_V = (3/2)nR$ . To find the entropy we use  $(\partial S/\partial T)_V = C_V/T = 3nR/2T$ , and also  $(\partial S/\partial V)_T = (\partial P/\partial T)_V = nR/V$ , so that

$$S = \frac{3}{2}nR \ln(T/T_0) + nR \ln(V/V_0) + ns_0$$

as per Eq. (6-5). To find the Helmholtz function, and thence the other potentials, we use  $(\partial F/\partial T)_V = -S$  and  $(\partial F/\partial V)_T = -P$ , so that

$$F = \frac{3}{2}nRT - \frac{3}{2}nRT \ln(T/T_0) - ns_0T - nRT \ln(V/V_0)$$

$$U = F + TS = \frac{3}{2}nRT = \frac{3}{2}nRT_0 (V_0/V)^{2/3} \exp[2(S - ns_0)/3nR]$$

$$H = U + PV = \frac{5}{2}nRT = \frac{5}{2}nRT_0 (T_0 P / P_0 T)^{2/3} \exp[2(S - ns_0)/3nR] \quad (8-21)$$

$$G = F + PV = \left(\frac{5}{2}nR - ns_0\right)T - \frac{3}{2}nRT \ln(T/T_0) + nRT \ln(V/V_0)$$

$$= n\mu = \left(\frac{5}{2}nR - ns_0\right)T - \frac{5}{2}nRT \ln(T/T_0) + nRT \ln(P/P_0)$$

$$\Omega = F - n\mu = \left(\frac{3}{2}nR - ns_0\right)T - \frac{3}{2}nRT \ln(T/T_0) - nRT \ln(V/V_0) - n\mu$$

where we have utilized the fact that  $U = 0$  when  $T = 0$  to fix the value of the constant of integration  $F_0$ . These can be checked with Eqs. (8-1) to (8-15) for self-consistency.

We note that the chemical potential per mole of the gas is a function of the intensive variables  $T$  and  $P$  only and is thus independent of  $n$  (as it must be). In fact we can see that all the thermodynamic potentials are  $n$  times a function which is independent of  $n$ , for  $V/V_0$  and  $T/T_0$  are both independent of  $n$ . This general property of thermodynamic potentials must be true for any system involving a single component (why?).

# 9

## Changes of Phase

Every substance (except helium) is a solid at sufficiently low temperatures. When it is in equilibrium it has a crystalline structure (glasses are supercooled liquids, not in equilibrium). As has already been mentioned and will be discussed later in some detail, entropy is a measure of randomness, so it will not be surprising to find that the entropy of a perfect crystal vanishes at absolute zero. (This statement is sometimes called the third law of thermodynamics.)

### The Solid State

To see how the entropy and the thermodynamic potentials of a solid change as the temperature is increased, we shall utilize the simplified equation of state of Eq. (3-6),

$$V = V_0(1 + \beta T - \kappa P); \quad P = (\beta/\kappa)T - [(V - V_0)/\kappa V_0] \quad (9-1)$$

Its heat capacities go to zero at zero temperature (see Fig. 3-1) and  $C_V$  rises to  $3nR$  at high temperatures. The capacity  $C_p$  is related to  $C_V$  according to Eq. (7-3). A simple formula that meets these requirements is

$$C_V = [3nRT^2/(\theta^2 + T^2)]; \quad C_p = C_V + (\beta^2 V_0/\kappa)T \quad (9-2)$$

The formulas do not fit too well near  $T = 0$ , but they have the right general shape and can be integrated easily. Constant  $\theta$  usually has a value less than  $100^\circ\text{K}$  so that by room temperature  $C_V$  is practically equal to  $3nR$ .

Using Eqs. (7-4) we compute the entropy as a function of  $T$  and  $P$ , remembering that  $S = 0$  when  $T = 0$  and  $P = 0$ ,

$$S = \frac{3}{2}nR \ln[1 + (T^2/\theta^2)] + (\beta^2 V_0/\kappa)T - \beta V_0 P \quad (9-3)$$

This formula makes it appear that  $S$  can become negative at very low temperatures if  $P$  is made large enough. Actually  $\beta$  becomes zero at  $T = 0$ , so that at very low temperatures  $S$  is independent of pressure. At moderate temperatures and pressures Eq. (9-3) is valid; curves of  $S$  as function of  $T$  for different values of  $P$  are shown in Fig. 9-1.

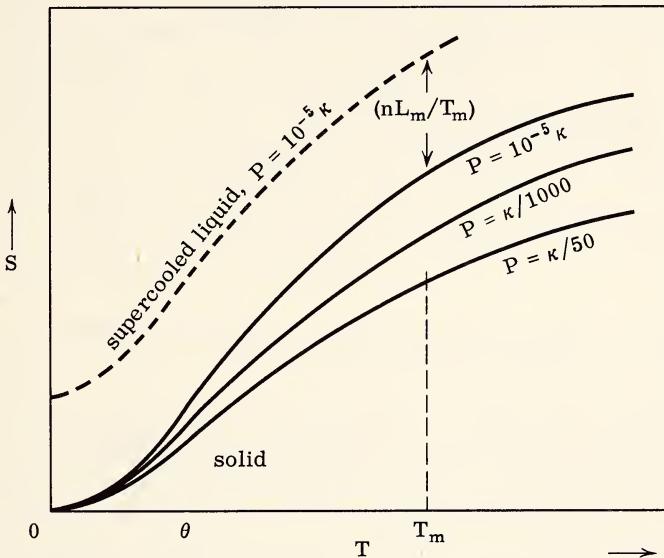


FIG. 9-1. Solid lines plot entropy of a solid as a function of  $T$  and  $P$ .

Next we use Eqs. (8-11) to compute the Gibbs function  $G$  as function of  $T$  and  $P$ ,

$$G = -\frac{3}{2}nRT \ln \left( 1 + \frac{T^2}{\theta^2} \right) + 3nRT - 3nR\theta \tan^{-1} (T/\theta) - (V_0\beta^2/2\kappa)T^2 + V_0P \left( 1 + \beta T - \frac{1}{2}\kappa P \right) + U_0 \quad (9-4)$$

where  $U_0$  is a constant of integration. Typical curves for this function are plotted in Fig. 9-2 for different values of  $P$ . We note that  $G$  is nearly constant at low temperatures, dropping somewhat at higher temperatures, as indicated by Eq. (8-20).

### Melting

If we add more and more heat (quasistatically) to the crystalline solid, holding the pressure constant at some moderate value, its tem-

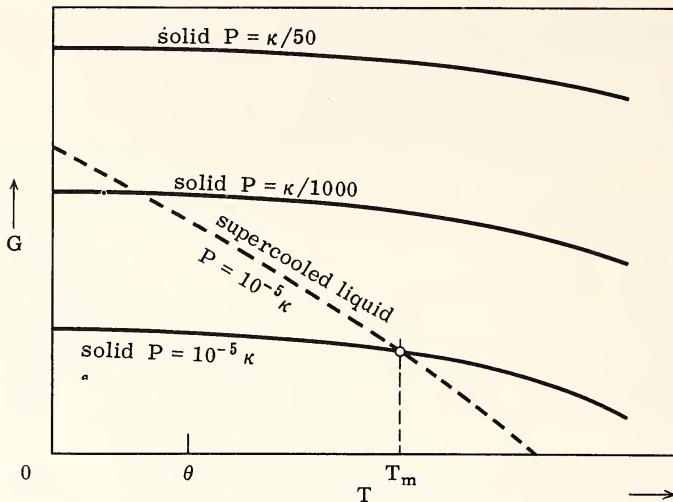


FIG. 9-2. Solid lines plot Gibbs function of a solid as a function of  $T$  and  $P$ .

perature rises until finally it melts, turning into a liquid with none of the regularities of the crystal. During the melting, addition of more heat simply melts more crystal; the temperature does not rise again until all is melted. The temperature  $T_m$  at which the melting occurs depends on the pressure, and the amount of heat required to melt 1 mole of the crystal  $L_m$  is called the *latent heat of melting* (it also is a function of the pressure).

We wish to ask how thermodynamics explains these facts, next to find whether it can predict anything about the dependence of  $T_m$ , the temperature of melting, on  $P$ , and of the latent heat of melting  $L_m$  on any of the thermodynamic quantities.

The answer to the first question lies in the discussion of Eq. (8-11), defining the Gibbs function. At any instant during the quasistatic process of heat addition, the temperature and pressure are constant; thus the material takes up the configuration which has the lowest value of  $G$  for that  $T$  and  $P$ . Below  $T_m$  the  $G$  for the solid is less than the  $G$  for the liquid; above  $T_m$  the liquid phase has the lower value of  $G$ ; if the process is carried out reversibly, all the material must melt at the temperature  $T_m(P)$ , at which the two  $G$ 's are equal (for the pressure  $P$ ).

If we could supercool the liquid to  $T = 0$  we would find its entropy to be larger than zero, because a liquid is irregular in structure. Furthermore the entropy of the liquid increases more rapidly with temperature than does the  $S$  for the solid. Since  $(\partial G / \partial T)_P = -S$ , even though at  $T = 0$  the  $G$  for the liquid is greater than the  $G$  for the

solid, it will drop more rapidly as  $T$  increases (see the dotted line of Fig. 9-2) until, at  $T = T_m$  the two  $G$ 's are equal; above  $T_m$  the liquid has the lower  $G$  and is thus the stable phase. Thus thermodynamics explains the sudden and complete change of phase at  $T_m$ .

The second question, raised earlier, is partly answered by pointing out that heat must be added to melt the material and that, at constant  $T$  an addition of heat  $dQ$  corresponds to an increase in entropy of the substance by an amount  $dQ/T$ . Thus the  $n$  moles of liquid, at the melting point  $T_m$ , has an entropy  $nL_m/T_m$  greater than the solid at the melting point, where  $nL_m$  is the heat required to melt  $n$  moles of the solid (at the specified  $P$ ). Thus a measurement of the latent heat of melting  $L_m$  enables one to compute the entropy of the liquid at  $T_m$ , in terms of the entropy of the solid (which is computed by integration from  $T = 0$ ), and further integration enables one to compute  $S$  for the liquid, for  $T$  greater than  $T_m$ , knowing the heat capacity and the equation of state of the liquid.

### Clausius-Clapeyron Equation

To answer the rest of the second question we utilize the fact that at the melting point the Gibbs function  $G_s(T_m, P)$  for the solid equals the Gibbs function  $G_l(T_m, P)$  for the liquid, no matter what the pressure  $P$ . In other words, as we change the pressure, the temperature of melting  $T_m(P)$ , the Gibbs function  $G_s$  for the solid and  $G_l$  for the liquid all change, but the change in  $G_s$  must equal the change in  $G_l$ , in order that the two  $G$ 's remain equal at the new pressure. Referring to Eq. (8-11) we see that this means that

$$dG_s = -S_s dT + V_s dP = dG_l = -S_l dT + V_l dP$$

or

$$(V_l - V_s) dP = (S_l - S_s) dT_m = (nL_m/T_m) dT_m$$

since the difference in entropy between a mole of liquid and a mole of solid is equal to the latent heat divided by the temperature of melting. Thus the equation relating  $T_m$  to  $P$  is

$$dT_m/dP = (T_m/nL_m)(V_l - V_s) \quad (9-5)$$

which is the *Clausius-Clapeyron equation*.

If the volume of the liquid is greater than the volume of the solid, then an increase of pressure will raise the temperature of melting so that, for example, if such a liquid is just above its melting point, an increase in pressure can cause it to solidify. Vice versa, if the solid is less dense than the liquid (as is the case with ice) an increase of

pressure lowers the melting point and pressure can cause such a solid, just below its melting point, to melt. Thus the fact that ice floats is related to the fact that ice skating is possible; ice skates ride on a film of water which has been liquefied by the pressure of the skate. In general, since  $V_1$  differs but little from  $V_s$ , the pressure must be changed by several thousand atmospheres to change  $T_m$  by as much as 10 per cent.

### Evaporation

If now heat is added to the liquid, its temperature will increase until another phase change occurs—the liquid *evaporates*. Here again the temperature remains constant at the temperature of vaporization  $T_v$  until all the liquid is converted into vapor. To be sure we understand what has taken place let us examine the process in more detail. We have tacitly assumed that the substance, first solid and then liquid, is confined in a container which adjusts its volume  $V$  so that it exerts a pressure  $P$  on *all parts* of the outer surface of the material. In other words we have assumed that the volume  $V$  is completely filled by the substance.

This may be difficult to do for the solid, but it is not hard to arrange it for the liquid. We provide the container with a piston, which exerts a constant force on the liquid and which can move to allow the liquid to expand at constant pressure  $P$ , and we make sure the liquid completely fills the container. In this case the liquid will stay liquid while we add heat, until its temperature reaches  $T_v(P)$ , when it must all be converted into gas (at a much greater volume but at the same pressure) before additional heat will raise the temperature beyond  $T_v$ . The temperature of vaporization  $T_v(P)$  is related to the pressure by another Clausius-Clapeyron equation,

$$dT_v/dP = (T_v/nL_v)(V_g - V_1) \quad (9-6)$$

where here  $L_v$  is the latent heat of evaporation per mole of the material (at pressure  $P$ ),  $V_1$  is the volume of the material as a liquid before evaporation, and  $V_g$  is its volume as a gas, after evaporation, at  $T_v$  and  $P$ . Since  $V_g$  is very much larger than  $V_1$ ,  $T_v$  changes much more rapidly with  $P$  than does  $T_m$ .

However our usual experience is not with the behavior of liquids that entirely fill a container, but with evaporation from the free surface of a liquid. When a liquid (or solid) does not completely fill a container, some of the substance evaporates into the free space until there is enough vapor there so that equilibrium between evaporation and condensation is reached. This equilibrium is only reached when the temperature of the liquid and vapor is related to the pressure of

the vapor in the free space above the liquid by the functional relationship we have been writing,  $T_v(P)$ , determined by Eq. (9-6). In the case we are now discussing it is better to reverse the functional relationship and write that the *vapor pressure*  $P_v$  is a function of the temperature  $T$ , and that the equation specifying this relationship is the reciprocal of (9-6),

$$\frac{dP_v}{dT} = [n L_v/T(V_g - V_1)] \quad (9-7)$$

This is the more-familiar form of the Clausius-Clapeyron equation.

The presence of another kind of gas in the space above the free surface of a liquid (or solid) only has an indirect effect on the amount of vapor present. The total pressure  $P$  on the liquid is now the sum of the partial pressures  $P_f$  of the foreign gas and  $P_v$  of the vapor. An addition of enough more of the foreign gas to increase this total pressure by  $dP$  (keeping  $T$  constant) will increase the Gibbs function of the liquid by  $dG_l = V_1 dP$  [see Eq. (8-11);  $dT = 0$ ], but the Gibbs function of the same amount of material in gaseous form is not affected by the foreign gas, so  $dG_g = V_g dP_v$ . For liquid and vapor to remain in equilibrium  $dG_l$  must equal  $dG_g$ ; consequently the relation between vapor pressure  $P_v(P, T)$  in the presence of a foreign gas and the total pressure  $P$  is given by

$$\frac{dP_v}{dP} = V_1/V_g$$

which may be integrated from the initial state where no foreign gas is present [ $P = P_v$  and  $P_v$  is the solution of Eq. (9-6)] to the final state where  $P = P_f + P_v$ . Since  $V_g$  is so much larger than  $V_1$ ,  $P_v$  changes very little as the foreign gas is added, but what change there is, is positive. Addition of foreign gas squeezes a little more vapor out of the liquid, rather than pushing some vapor back into the liquid.

Water in a dish open to the air is not in equilibrium unless the partial pressure of water vapor in the air happens to be exactly equal to the vapor pressure  $P_v(T)$  for the common temperature  $T$  of air and water (this is the condition of 100 per cent humidity). If the common temperature is above this, the water continues to evaporate until it is all gone, the evaporation proceeding more rapidly as the temperature is raised, until the boiling point is reached, when  $P_v(T)$  is equal to the total atmospheric pressure; the gas immediately above the water is all water vapor, and the water boils rapidly away.

The latent heats of evaporation are usually 10 to 50 times greater than the corresponding latent heats of melting, corresponding to the fact that it takes much more work to pull the material into a tenuous vapor than it does to change it into a liquid, which disrupts the crystal structure but doesn't pull the atoms much further apart.

At very high pressures there are also phase changes in the solid state; the crystal structure of the solid changes, with accompanying latent heat, change of volume, and relationship between  $P$  and  $T$  for the change given by an equation such as (9-6).

### Triple Point and Critical Point

We have just seen that the melting temperature is nearly independent of pressure, whereas the temperature of vaporization is strongly dependent on  $P$ . Therefore as  $P$  decreases, the two curves, one for  $T_v$ , the other for  $T_m$ , converge. This is shown in Fig. 9-3, where the

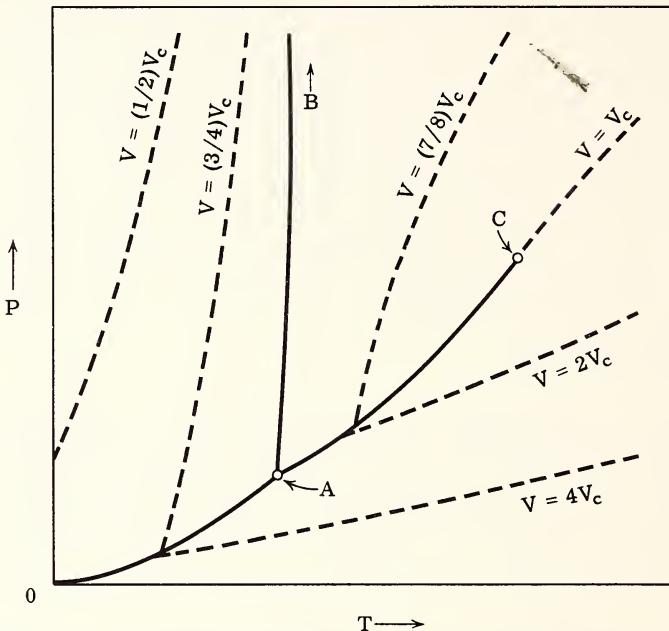


FIG. 9-3. Phase diagram for a material that expands upon melting. Solid lines are the curves for phase change, dashed lines those for constant volume.

curve AB is the melting-point curve and AC that for vaporization. The two meet at the *triple point* A, which is the only point where solid, liquid, and vapor can coexist in equilibrium. Below this pressure the liquid is not a stable phase and along the curve 0A the solid transforms directly into the vapor (sublimation). The shape of curve 0A is governed by an equation similar to (9-6), with a latent heat of sublimation  $L_s$  (equal to  $L_m + L_v$  at the triple point). The dashed lines of Fig. 9-3 are lines of  $P$  against  $T$  for different values of  $V$ , intersections of the PVT surface by planes parallel to the PT plane.

As the pressure is increased, keeping  $T = T_v(P)$  so that we follow the curve AC, the differences  $(V_g - V_l)$  and  $(S_g - S_l) = (L_v/T_v)$  between gas and liquid diminish until at C, the *critical point*, there ceases to be any distinction between liquid and gas and the curve AC terminates. There seems to be no such termination of the curve AB for melting; the difference between the regularly structured solid and the irregular liquid remains for pressures up to the maximum so far attained; it may be that curve AB continues to infinity.

The PT plane is only one way of viewing the PVT surface representing the equation of state. Another sometimes more useful projection is the one on the PV plane. In Fig. 9-4 are plotted the dashed curves of P against V for different values of T, corresponding to

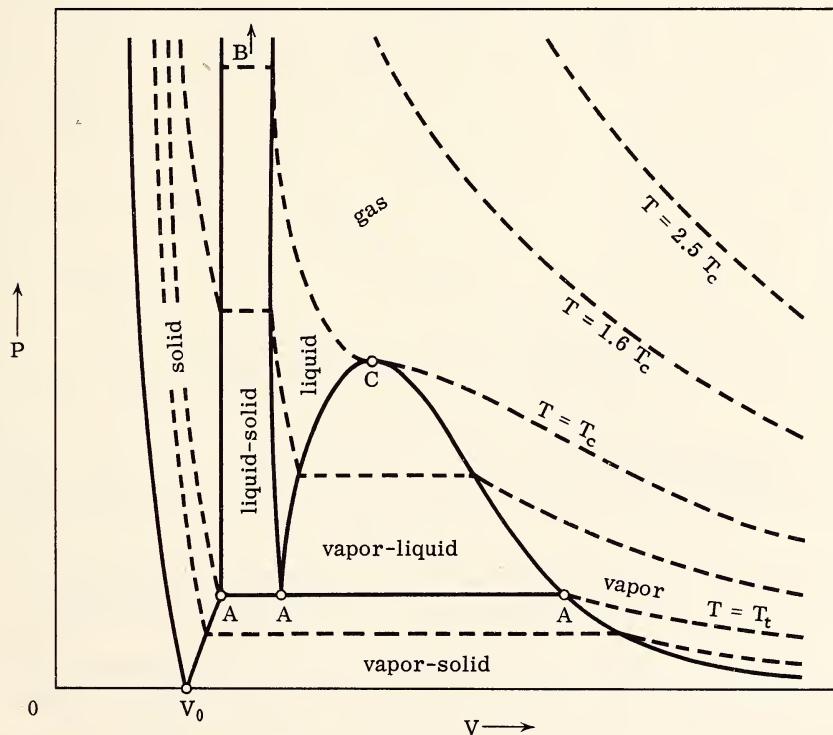


FIG. 9-4. PV curves (dashed lines) for the material of Fig. 9-3. Solid lines are projections on the PV plane of the solid lines of Fig. 9-3.

intersections of the PVT surface with planes of constant  $T$ , parallel to the PT plane. The regions where the PV curves are horizontal are where there is a phase change. The boundaries of these regions, ABA and ACA, projected on the PT plane, are the curves AB and

AC of Fig. 9-3. It is clearer from Fig. 9-4, why C is the critical point. The line AAA corresponds to the triple point A of Fig. 9-3.

The SPT surface is also divided into the various phase regions. Figure 9-5 shows a part of this surface, projected on the ST plane, the surface being ruled with the dashed lines of constant pressure.

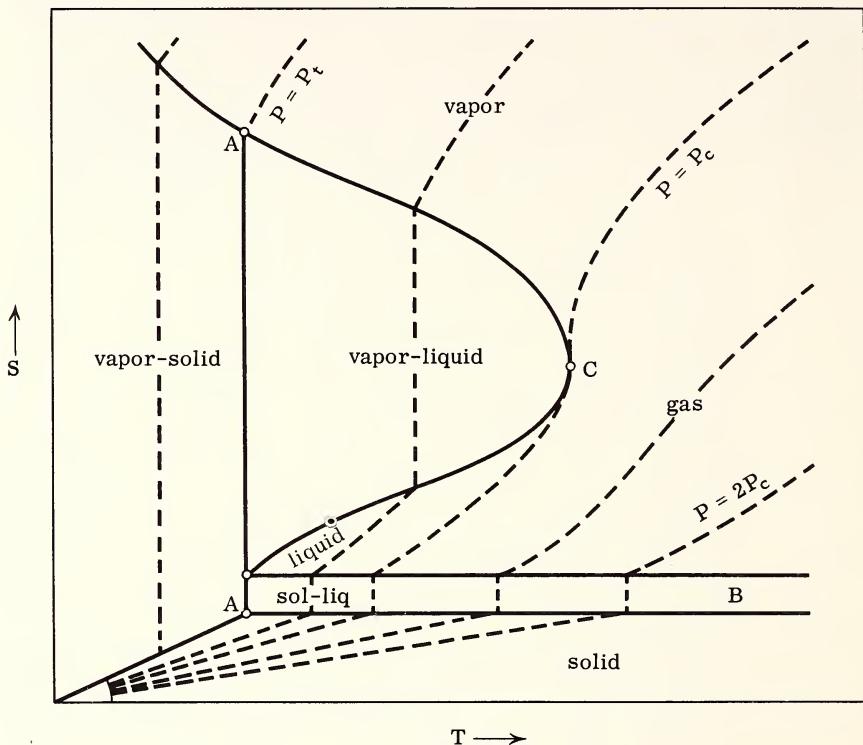


FIG. 9-5. Entropy as a function of T (dashed curves) for various values of  $P$ , for the material of Fig. 9-3.

We see that, as pressure is kept constant and  $T$  is increased,  $S$  increases steadily until a phase change occurs, when  $S$  takes a sudden jump, of amount  $L/T$ , and then continues its steady increase with temperature in the new phase. Entropy change is largest between solid and vapor, not because the latent heat is so much larger for this phase change, but because it occurs at low temperature and  $S_s - S_g = L_s/T_s$ , where  $T_s$  is small.

The GPT surface for the Gibbs function, projected on the GT plane, is shown in Fig. 9-6. The dashed lines correspond to the intersections of the surface with planes of constant  $P$ , parallel to the GT

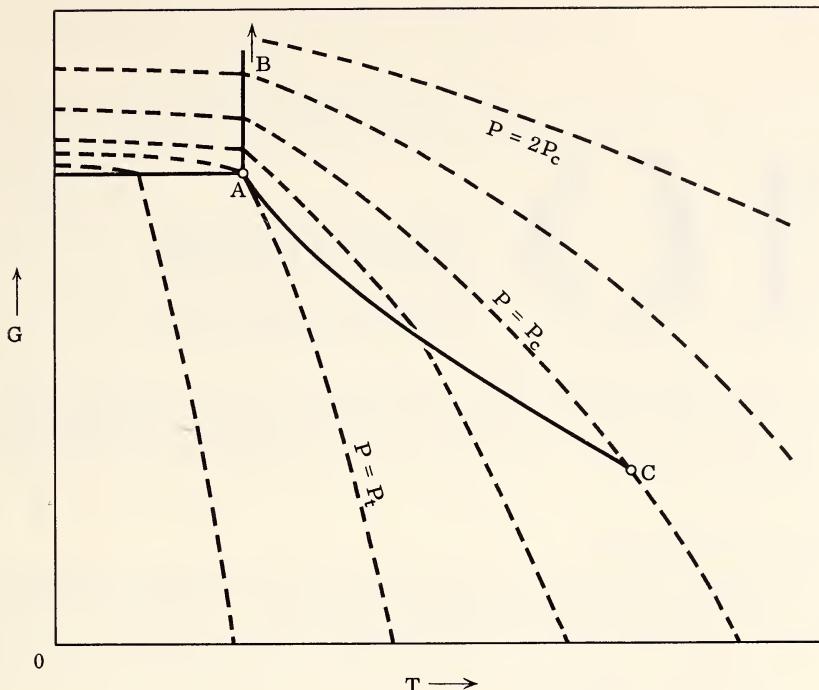


FIG. 9-6. Gibbs function versus temperature (dashed lines) for various values of  $P$ , for the material of Fig. 9-3.

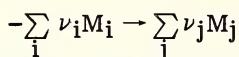
plane. The solid lines correspond to the phase changes. As noted earlier in this chapter,  $G$  does not change suddenly during a phase change, as do  $V$  and  $S$ ; only the slopes  $(\partial G / \partial P)_T = V$  and  $(\partial G / \partial T)_V = -S$  (the slopes of the dashed lines of Fig. 9-6) change discontinuously across the phase-change boundaries. By taking gradients the curves of Figs. 9-4 and 9-5 can be obtained from Fig. 9-6 or, vice versa, the curves of Fig. 9-6 can be obtained by integration of the data on curves in Figs. 9-4 and 9-5.

# 10 Chemical Reactions

The Gibbs function also is of importance in describing chemical processes. Since most chemical reactions take place at constant temperature and pressure, the reaction must go in the direction of decreasing  $G$ .

## Chemical Equations

A chemical reaction is usually described by an equation, such as  $2H_2 + O_2 \rightarrow 2H_2O$ , which we can generalize as



stating that a certain number,  $-\nu_i$ , of molecules of the *initial reactants*  $M_i$  will combine to produce a certain number,  $\nu_j$ , of molecules of *final products*  $M_j$ . Of course the chemical reaction can run in either direction, depending on the circumstances; we have to pick a direction to call positive. This is arbitrarily chosen to be the direction in which the reaction generates heat. We then, also arbitrarily, place all the terms in the equation on the right, so that it reads



In this form the  $\nu$ 's which are negative represent initial reactants and those which are positive represent final products ( $\nu$  for the  $H_2$  in the example would be  $-2$ ,  $\nu$  for the  $O_2$  would be  $-1$ , and  $\nu$  for the  $H_2O$  would be  $+2$ ). The number  $\nu_i$  is called the *stoichiometric coefficient* for  $M_i$  in the reaction.

For a chemical reaction to take place, more than one sort of material must be present either initially or finally, and the numbers  $n_i$  of moles of the reacting materials will change during the reaction. Referring to Eqs. (8-11) and (8-17) we see that the Gibbs function and its change during the reaction are

$$G = \sum_i n_i \mu_i; \quad dG = -S dT + V dP + \sum_i \mu_i dn_i \quad (10-2)$$

where  $\mu_i$  is the chemical potential of the  $i$ -th component of the reaction, which we see is the Gibbs function per mole of the material  $M_i$ . We shall see shortly how it depends on  $T$  and  $P$  and how it can be measured.

A chemical reaction of the sort described by Eq. (10-1) produces a change in the  $n$ 's in an interrelated way. If  $dn_1 = \nu_1 dx$  moles of material  $M_1$  appear during a given interval of time while the reaction progresses, then  $dn_i = \nu_i dx$  moles of material  $M_i$  will appear during the same interval of time (or will disappear, if  $\nu_i$  is negative). For example, if  $2dx$  moles of  $H_2O$  appear, simultaneously  $2dx$  moles of  $H_2$  and  $dx$  moles of  $O_2$  will disappear (i.e.,  $dn_H = -2dx$  and  $dn_O = -dx$ ). In other words, during a chemical reaction at constant  $T$  and  $P$ , the change in the Gibbs function is

$$dG = \sum_i \mu_i \nu_i dx$$

In accordance with the discussion of Eq. (8-11) we should expect the reaction at constant  $T$  and  $P$  to continue spontaneously, with consequent decrease of  $G$ , until  $G$  becomes minimum at equilibrium. Thus, at equilibrium, at constant  $T$  and  $P$ ,  $dG/dx = 0$ , or

$$\sum_i \mu_i \nu_i = 0 \quad \text{for equilibrium} \quad (10-3)$$

At equilibrium the relative proportions of the reactants and the products must adjust themselves so that the  $\mu$ 's, which are functions of  $T$ ,  $P$ , and the concentration of the  $i$ -th component, satisfy Eq. (10-3).

### Heat Evolved by the Reaction

During the progress of the reaction, if carried out at constant  $T$  and  $P$ , any evolution of heat would be measured as a change in enthalpy since, as we have already remarked [see the discussion after Eq. (8-7)], enthalpy is the heat content of the system at constant pressure. But since, from Eq. (8-17),

$$H = G + TS = G - T(\partial G / \partial T)_{pn} \quad (10-4)$$

the change in  $H$  as the parameter  $x$  changes at constant  $T$  and  $P$  is

$$dH = \left( \frac{dG}{dx} \right) dx - T \left[ \frac{\partial}{\partial T} \left( \frac{dG}{dx} \right) \right]_{pn} dx = \left[ \sum \mu_i \nu_i - T \left( \frac{\partial}{\partial T} \sum \mu_i \nu_i \right)_{pn} \right] dx$$

At equilibrium  $\sum \mu_i \nu_i = 0$  and the reaction ceases. If we measure the heat evolved per change  $dx$  when the system is close to equilibrium, the rate of evolution of heat becomes

$$dH/dx = -T \left( \frac{\partial}{\partial T} \sum \mu_i \nu_i \right)_{pn} \quad (10-5)$$

Thus by measuring the rates of change with temperature,  $(\partial \mu_i / \partial T)_{pn}$ , of the chemical potentials of the substances involved, we can predict the amount of heat evolved during the reaction. Or, vice versa, if we can measure (or can compute, by quantum mechanics) the heat evolved when  $\nu_i$  moles of substance  $M_i$  appear or disappear, by dissociating into their constituent atoms or by reassociating the atoms into the product molecules, we can predict the rate of change of the  $\mu$ 's with temperature.

### Reactions in Gases

To see how all this works out, we take the simple case of a mixture at high-enough temperature so that all the components are perfect gases. The number of moles of the  $i$ -th component is  $n_i$  and the total number of moles in the mixture is  $n = \sum n_i$ . The relative proportions of the different gases can be given in terms of their *concentrations*  $\chi_i = n_i/n$ , so that  $\sum \chi_i = 1$ , or they can be expressed in terms of their *partial pressures*  $p_i = \chi_i P$ . Each mode of expression has its advantages.

For example, we can compute the Gibbs function per mole  $\mu_i$  of  $M_i$ , as a perfect gas, in terms of  $T$  and  $p_i$ . Using the procedures of Chapter 8 [see Eqs. (8-21) for example], we have, for a perfect gas for which  $C_p(T) = C_v(T) + nR$ ,

$$(\partial S / \partial T)_p = (1/T)C_p(T); \quad (\partial S / \partial P)_T = -(\partial V / \partial T)_p = -(nR/P)$$

so

$$S = S_0 + \int_{T_0}^T C_p(T)(dT/T) - nR \ln(P/P_0)$$

and, since

$$(\partial G / \partial T)_p = -S \quad \text{and} \quad (\partial G / \partial P)_T = V = nRT/P$$

we have

$$\partial G / \partial n_i = \mu_i = RT \ln(P_i/P_0) + g_i(T) = RT[\ln(P/P_0) + \ln \chi_i] + g_i(T) \quad (10-6)$$

where

$$g_i(T) = g_i(T_0) - s_0(T - T_0) - \int_{T_0}^T dT \int_{T_0}^T c_p(T) (dT/T)$$

$s_0$  being the entropy per mole of the  $i$ -th component at  $T = T_0$  and  $c_p$  being the specific heat, the heat capacity per mole at constant pressure of  $M_i$  in its gaseous form.

The equation for chemical equilibrium (10-3) then takes on the form

$$RT \sum_i \nu_i \left[ \ln \left( \frac{P}{P_0} \right) + \ln \chi_i + \frac{g_i}{RT} \right] = 0$$

or

$$\Pi \chi_i^{\nu_i} = (P_0/P)^{\sum \nu_i} K(T); \quad K(T) = \exp \left[ -\sum \nu_i g_i(T)/RT \right] \quad (10-7)$$

where the sign  $\Pi$  indicates the product of all the terms  $\chi_i^{\nu_i}$  for all values of  $i$ .

The quantity  $K(T)$  is called the *equilibrium constant* and the equation determining the  $\chi$ 's at equilibrium is called the *law of mass action*. To see how this goes, we return to the reaction between  $H_2$  and  $O_2$ . Suppose initially  $n_1$  moles of  $H_2$  and  $n_2$  moles of  $O_2$  were present and suppose during the reaction  $x$  moles of  $O_2$  and  $2x$  moles of  $H_2$  combined to form  $2x$  moles of  $H_2O$ . The total number of moles then present would be  $(n_1 - 2x) + (n_2 - x) + 2x = n_1 + n_2 - x$ , so the concentrations at equilibrium and the stoichiometric coefficients for the reaction are

$$\text{for } H_2: \quad \chi_1 = [(n_1 - 2x)/(n_1 + n_2 - x)] \quad \nu_1 = -2$$

$$\text{for } O_2: \quad \chi_2 = [(n_2 - x)/(n_1 + n_2 - x)] \quad \nu_2 = -1$$

$$\text{for } H_2O: \quad \chi_3 = [2x/(n_1 + n_2 - x)] \quad \nu_3 = 2$$

and the law of mass action becomes

$$\frac{2x(n_1 + n_2 - x)}{(n_1 - 2x)^2 (n_2 - x)} = \left( \frac{P}{P_0} \right) \exp \left[ \frac{2g_1 + g_2 - 2g_3}{RT} \right] = \left( \frac{P}{P_0} \right) K(T)$$

from which we can solve for  $x$ . Since this particular reaction is strongly exothermic,  $2g_1 + g_2$  is considerably larger than  $2g_3$  and the exponential  $K(T)$  is a very large quantity unless  $T$  is large. Consequently, for  $P \approx P_0$  and for moderate temperatures,  $x$  will be close to  $(1/2)n_1$  or  $n_2$ , whichever is smaller, i.e., the reaction will go almost to completion, using up nearly all the constituent in shorter supply. If there is a deficiency in hydrogen, for example, so that  $n_1 < 2n_2$ , then we set  $x = (1/2)n_1 - \delta$ , where  $\delta$  is small, and

$$(2\delta)^2 \simeq \frac{n_1[n_2 + (1/2)n_1]}{[n_2 - (1/2)n_1]} \left( \frac{P_0}{P} \right) \exp \left( \frac{2g_3 - 2g_1 - g_2}{RT} \right)$$

$$\chi_1 \simeq \frac{2\delta}{n_2 + (1/2)n_1}; \quad \chi_2 \simeq \frac{n_2 - (1/2)n_1}{n_2 + (1/2)n_1}; \quad \chi_3 \simeq \frac{n_1}{n_2 + (1/2)n_1}$$

the system coming to equilibrium with a very small amount,  $2\delta$  moles, of  $H_2$  left. This amount increases with decrease of pressure  $P$  and with increase of  $T$ .

To see how the equilibrium constant  $K(T)$  changes with temperature, we can utilize Eqs. (10-3), (10-5), and (10-6). We have

$$\begin{aligned} \frac{d}{dT} \ln K &= \frac{1}{RT^2} \left[ \sum \nu_i g_i - T \sum \nu_i g_i' \right] \\ &= \frac{1}{RT^2} \left[ \sum \nu_i \mu_i - T \sum \nu_i \left( R \ln \frac{P_i}{P_0} + g_i' \right) \right] \\ &= \frac{1}{RT^2} \left[ -T \left( \frac{\partial}{\partial T} \sum \nu_i \mu_i \right)_{pT} \right] = \frac{1}{RT^2} \left( \frac{\partial H}{\partial x} \right)_{Tp} \end{aligned} \quad (10-8)$$

where  $g_i' = (dg_i/dT)$ . The rate of change of the equilibrium constant  $K(T)$  with temperature is thus (for gas reactions) proportional to the amount of heat evolved  $(\partial H/\partial x)_{pT}$  per unit amount of reaction at equilibrium at constant  $T$  and  $P$ . This useful relationship is known as Van't Hoff's equation.

### Electrochemical Processes

Some chemical reactions can occur spontaneously in solution, generating heat, or they can be arranged to produce electrical energy instead of heat. For example, a solution of copper sulfate contains free copper ions in solution. The addition of a small amount,  $\Delta n$  moles, of metallic zinc in powder form will cause  $\Delta n$  moles of copper to appear in metallic form, the zinc going into solution, replacing the  $Cu$  ions. At the same time an amount  $W \Delta n$  of heat is released. If the reaction takes place at constant pressure and temperature, Eq. (8-7) indicates that  $W \Delta n$  must equal the difference in enthalpy between the initial state, with  $Cu$  in solution, and the final state, with  $Zn$  in solution,

$$H_1 - H_2 = W \Delta n \quad (10-9)$$

However the same reaction can take place in a battery, having one electrode of copper in a  $CuSO_4$  solution and the other electrode (the negative pole of the battery) of zinc, surrounded by a  $ZnSO_4$  solution,

the two solutions being in contact electrically and thermally. If the battery now discharges an amount  $\Delta C$  coulombs of charge through a resistor, or a motor to produce work, more Cu will be deposited on the Cu electrode and an equal number of moles,  $\Delta n$ , of Zn will leave the zinc electrode to go into solution. In this case the energy of the reaction goes into electromechanical work; for every charge  $\Delta C$  discharged by the battery,  $\varepsilon \Delta C$  joules of work are produced, where  $\varepsilon$  is the equilibrium voltage difference between the battery electrodes.

If the battery is kept at constant temperature and pressure during the quasistatic production of electrical work, Eq. (8-11)(which in this case can be written  $dG = -S dT + V dP + \varepsilon dC$ ) shows that the work done equals the change in the Gibbs function caused by the reaction. In other words  $\varepsilon \Delta C = \Delta G$ . But Eq. (10-4) shows the relationship between enthalpy and Gibbs function, which in this case can be written

$$G = H_1 - H_2 + T(\partial \Delta G / \partial T)_p \quad (10-10)$$

which, with Eq. (10-9), provides a relationship between the electrical properties of the battery and the thermal properties of the related chemical reaction.

If the ions have a valency  $z$  (for Zn and Cu,  $z = 2$ ) then a mole of ions possesses a charge  $z F$ , where  $F$  is the Faraday constant  $9.65 \times 10^7$  coulombs per mole. Thus the charge  $\Delta C$  is equal to  $z F \Delta n$ , where  $\Delta n$  is the number of moles of Zn that goes into solution (or the number of moles of Cu that is deposited). Combining Eqs. (10-9) and (10-10), we obtain the equation (using  $\Delta G = z F \varepsilon \Delta n$  and dividing by  $z F \Delta n$ )

$$\varepsilon = (W/zF) + T(\partial \varepsilon / \partial T)_p \quad (10-11)$$

relating the emf of the cell and the change in emf with temperature to the heat  $W$  evolved in the corresponding chemical reaction; and, by use of Eq. (10-5), we have derived a means of obtaining empirical values of the chemical potentials  $\mu_i$ . Thus electrical measurements, which can be made quite accurately, can be used instead of thermal measurements to measure heats of reaction. Equation (10-11) is called the Gibbs-Helmholtz equation.



# II

## KINETIC THEORY



# 11 Probability and Distribution Functions

We have now sketched out the main logical features of thermodynamics and have discussed a few of its applications. We could easily devote the rest of this text to other applications, covering all branches of science and engineering. But, as physicists, it is more appropriate for us to go on to investigate the connection between the thermal properties of matter in bulk and the detailed properties of the atoms that constitute this matter. The connection, as we saw in Chapter 2, must be a statistical one and thus will be expressed in terms of probabilities.

## Probability

A basic concept is hard to define, except circularly. Probability is in part subjective, a quantization of our expectation of the outcome of some event (or trial) and only measurable if the event or trial can be repeated several times. Suppose one of the possible outcomes of a trial is A. We say that the probability that the trial results in A is  $P(A)$  if we expect that, out of series of  $N$  similar trials, roughly  $NP(A)$  of them will result in A. We expect that the fraction of the trials which do result in A will approach  $P(A)$  as the number of trials increases. A standard example is the gambler's six-sided die; a 5 doesn't come up regularly every sixth time the die is thrown, but if a 5 comes up 23 times in 60 throws and also 65 times in the following 240 throws, we begin seriously to doubt the symmetry of the die and/or the honesty of the thrower.

If result A occurs for every trial, then  $P(A) = 1$ . If other events sometimes occur, such as event B, then the probability that A does not occur in a trial,  $1 - P(A)$ , is not zero. It may be possible that both A and B can occur in a single trial; the probability of this happening is written  $P(AB)$ . Simple logic will show that the probability of either A or B or both occurring in a trial is

$$P(A + B) = P(A) + P(B) - P(AB) \quad (11-1)$$

Relationships between probabilities are often expressed in terms of the *conditional probabilities*  $P(A|B)$  that A occurs in a trial *if* B also occurs, and  $P(B|A)$  the probability that B occurs *if* A occurs as well. We can see that

$$P(AB) = P(A|B) P(B) = P(B|A) P(A) = P(BA) \quad (11-2)$$

A simple example is in the dealing of well-shuffled cards. The chance of a heart being dealt is  $P(H) = 1/4$ , the chance that the card is a seven *if* it is a heart is  $P(7|H) = 1/13$  and therefore the probability that the card dealt is the seven of hearts is  $P(7H) = P(7|H) P(H) = (1/13)(1/4) = 1/52$ .

If the probability of A occurring in a trial is not influenced by the simultaneous presence or absence of B, i.e., if  $P(A|B) = P(A)$ , then A and B are said to be *independent*. When this is the case,

$$\begin{aligned} P(A|B) &= P(A); & P(B|A) &= P(B); & P(AB) &= P(A) P(B) \\ [1 - P(A + B)] &= [1 - P(A)][1 - P(B)] \end{aligned} \quad (11-3)$$

Saying it in words, if A and B are independent, then the chance of *both* A and B occurring in a trial is the *product*  $P(A) P(B)$  of their separate probabilities of occurrence and the probability  $1 - P(A + B)$  that *neither* A nor B occur is the product of their separate probabilities of nonoccurrence. In the example of the dealing of cards, since the chance  $P(7)$  of a seven being dealt is the same as the conditional probability  $P(7|H)$  that a seven is dealt *if* it is a heart, the occurrence of a seven is independent of the occurrence of a heart. Thus the probability that the card dealt is a seven of hearts is  $P(7) P(H) = (1/13)(1/4) = 1/52$ , and the chance that it is neither a seven nor a heart is  $(12/13)(3/4) = 9/13$ .

If the trial is such that, when A occurs B cannot occur and vice versa, then A and B are said to be *exclusive* and

$$P(AB) = P(A|B) = P(B|A) = 0$$

so that

$$P(A + B) = P(A) + P(B) \quad (11-4)$$

The chance of *either* A or B occurring, when A and B are exclusive, is thus the *sum* of their separate probabilities of occurrence. For example, the result that a thrown die comes up a 5 is exclusive of its

coming up a 1; therefore the chance of either 1 or 5 coming up is  $(1/6) + (1/6) = 1/3$ .

Probabilities are useful in discussing *random* events. The definition of randomness is as roundabout as the definition of probability. The results of successive trials are randomly distributed if there is no pattern in the successive outcomes, if the only prediction we can make about the outcome of the next trial is to state the probabilities of the various outcomes.

### Binomial Distribution

The enumeration of the probabilities of all the possible outcomes of an event is called a *distribution function*, or a probability distribution. For example, suppose the "event" consists of  $N$  independent and equivalent trials, such as the throwing of  $N$  dice (or the throwing of one die  $N$  times). If the probability of "success" in one trial is  $p$  (if "success" is for the die to come up a 1, for example, then  $p = 1/6$ ), it is not difficult to show that the probability of exactly  $n$  successes in  $N$  trials (not distinguishing the order of failures and successes) is

$$P_n(N) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n} \quad (11-5)$$

All possible results of the event are included in the set of values of  $n$  from 0 to  $N$ , so the set of probabilities  $P_n(N)$  is a distribution function; this particular one is called the *binomial distribution*.

When the possible results of the event are denumerable, as they are for the binomial distribution, the individual probabilities can be written  $P_n$ , where  $n$  is the integer labeling one of the possible results. If the results are mutually exclusive,

$$\sum_n P_n = 1 \quad (11-6)$$

where the sum is taken over the values of  $n$  corresponding to all possible results. For example, for the binomial distribution,

$$P_n(N) = \sum_{n=0}^N \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n} = [p + (1-p)]^N = 1$$

Suppose the *value* of result  $n$  is  $x(n)$ . The *expected value*  $\langle x \rangle$  of an event is then the weighted average

$$\langle x \rangle = \sum_n x(n) P_n \quad (11-7)$$

Any individual event would not necessarily result in this value, but we should expect that the average value of a series of similar events would tend to approach  $\langle x \rangle$ . A measure of the variability of individual events is the *variance* ( $\sigma_x$ )<sup>2</sup> of  $x$ , the mean square of the difference between the actual value of an event and its expected value,

$$(\sigma_x)^2 = \sum_n [x(n) - \langle x \rangle]^2 P_n = \langle x^2 \rangle - \langle x \rangle^2 \quad (11-8)$$

The square root of the variance  $\sigma_x$  is called the *standard deviation* of  $x$  for the particular distribution.

### Random Walk

To make these generalities more specific let us consider the process called the *random walk*. We imagine a particle in irregular motion along a line; at the end of each period of time  $\tau$  it either has moved a distance  $\delta$  to the right or a distance  $\delta$  to the left of where it was at the beginning of the period. Suppose the direction of each successive "step" is independent of the direction of the previous one, that the probability that the step is in the positive direction is  $p$  and the probability that it is in the negative direction is  $1 - p$ . Then the probability that during  $N$  periods the particle has made  $n$  positive steps, and thus  $N - n$  negative steps, is the binomial probability  $P_n(N)$  of Eq. (11-5).

The net displacement after these  $N$  periods,  $x(n) = (2n - N)\delta$ , might be called the "value" of the  $N$  random steps of the particle. The expected value of this displacement after  $N$  steps can be computed in terms of the expected value  $\langle n \rangle$  of  $n$ ,

$$\langle n \rangle = \sum_{n=0}^N n P_n(N) = p \sum_{k=0}^{N-1} \frac{N!}{k!(N-1-k)!} p^k (1-p)^{N-1-k} = pN$$

so

$$\langle x(n) \rangle = \langle (2n - N)\delta \rangle = (2p - 1)N\delta; \quad k = n - 1 \quad (11-9)$$

When  $p = 1$  (the particle always moves to the right) the expected displacement after  $N$  steps is  $N\delta$ ; when  $p = 0$  (the particle always moves to the left) it is  $-N\delta$ ; when  $p = 1/2$  (the particle is equally likely to move right as left at each step) the expected displacement is zero. The variability of the actual result of any particular set of  $N$  steps can be obtained from  $\langle n \rangle$  and  $\langle n^2 \rangle$ ,

$$\langle n^2 \rangle = \langle n(n-1) \rangle + \langle n \rangle = p^2 N(N-1) + pN$$

so that

$$\begin{aligned} (\sigma_x)^2 &= \langle [x(n) - N(2p - 1)\delta]^2 \rangle = \langle (2n - 2pN)^2 \delta^2 \rangle \\ &= 4Np(1 - p)\delta^2 \end{aligned} \quad (11-10)$$

When  $p = 1$  (all steps to right) or  $p = 0$  (all steps to left) the variance  $(\sigma_x)^2$  of the displacement is zero; when  $p = 1/2$  the variance is greatest, being  $N\delta^2$  (the standard deviation is then  $\sigma_x = \delta\sqrt{N}$ ).

### The Poisson Distribution

When the outcome of an event has a continuous range of values, the sums discussed heretofore must be changed to integrals and we must talk about a *probability density*  $f(x)$ , defined so that  $f(x) dx$  is the probability that the result lies between  $x$  and  $x + dx$ . The representative example is that of dots distributed at random along a line. The line may be the distance traveled by a molecule of a gas and the dots may represent collisions with other molecules, or the line may be the time axis and the dots may be the instants of emission of a gamma ray from a piece of radioactive material; in any of these cases a dot is equally likely to be found in any element  $dx$  of length of the line. If the mean density of dots is  $(1/\lambda)$  per unit length ( $\lambda$  being thus a mean distance between dots) then the chance that a randomly chosen element  $dx$  of the line will contain a dot is  $(dx/\lambda)$ , independent of the position of  $dx$ .

We can now compute the probability  $P_0(x/\lambda)$  that no dot will be found in an interval of length  $x$  of the line, by setting up a differential equation for  $P_0$ . The probability  $P_0[(x+dx)/\lambda]$  that no dots are in an interval of length  $x + dx$  is equal to the product of the probability  $P_0(x)$ , that no dots are in length  $x$ , times the probability  $1 - (dx/\lambda)$ , that no dot is in the additional length  $dx$  [see Eq. (11-3)]. Using Taylor's theorem we have

$$P_0\left(\frac{x+dx}{\lambda}\right) = P_0\left(\frac{x}{\lambda}\right) + \frac{d}{dx} P_0\left(\frac{x}{\lambda}\right) dx = P_0\left(\frac{x}{\lambda}\right) \left[1 - \left(\frac{dx}{\lambda}\right)\right]$$

or

$$\frac{d}{dx} P_0\left(\frac{x}{\lambda}\right) = -\left(\frac{1}{\lambda}\right) P_0\left(\frac{x}{\lambda}\right) \quad \text{so that} \quad P_0\left(\frac{x}{\lambda}\right) = e^{-x/\lambda} \quad (11-11)$$

The probability that no dot is included in an interval of length  $x$ , placed anywhere along the line, thus decreases exponentially as  $x$  increases, being unity for  $x = 0$  (it is certain that no dot is included in

a zero-length interval), being 0.368 for  $x = \lambda$ , the mean distance between dots (i.e., a bit more than a third of the intervals between dots are larger than  $\lambda$ ) and dropping off rapidly to zero for  $x$  larger than  $\lambda$ .

The probability density for this distribution is the derivative of  $P_0$ ; it supplies the answer to the following question. We start at an arbitrary point and move along the line: What is the probability that the *first* dot encountered lies between  $x$  and  $x + dx$  from the start? Equation (11-3) indicates that this probability,  $f(x) dx$ , is equal to the product of the probability  $e^{-x/\lambda}$  that no dot is encountered in length  $x$ , times the probability  $dx/\lambda$  that a dot *is* encountered in the next interval  $dx$ . Therefore

$$f(x) = (1/\lambda)e^{-x/\lambda} \quad (11-12)$$

is the probability density for encountering the first dot at  $x$  (i.e.,  $f dx$  is the probability that the first dot is between  $x$  and  $x + dx$ ).

As with discrete distributions, one can compute expected values and variances for probability densities. For example, the expected distance one goes, from an arbitrarily chosen point on the line, before a dot is encountered is

$$\langle x \rangle = \int_0^\infty x f(x) dx = \lambda \int_0^\infty u e^{-u} du = \lambda \quad (11-13)$$

for the randomly distributed dots. The variance of this distance is

$$(\sigma_x)^2 = \int_0^\infty (x - \lambda)^2 f(x) dx = \lambda^2 \quad (11-14)$$

so the standard deviation of  $x$ ,  $\sigma_x = \lambda$ , is as large as the mean value  $\lambda$  of  $x$ , an indication of the variability of the interval sizes between the randomly placed dots.

We can go on to ask what is the probability  $P_1(x/\lambda)$  of finding just one dot in an interval of length  $x$  of the line. This is obtained by using Eq. (11-3) again to show that the probability of finding the first dot between  $y$  and  $y + dy$  from the beginning and no other dot between this and the end of the interval  $x$  is equal to the product  $f(y) dy P_0[(x-y)/\lambda]$ . Then we can use Eq. (11-4) to show that the probability  $P_1(x/\lambda)$  that the one dot is somewhere within the interval  $x$  is the integral

$$P_1(x/\lambda) = \int_0^x (1/\lambda)e^{-y/\lambda} dy e^{-(x-y)/\lambda} = (x/\lambda) e^{-x/\lambda}$$

An extension of this argument quickly shows that the probability that there are exactly  $n$  dots in the interval  $x$  is

$$P_n(x/\lambda) = \int_0^x f(y) dy P_{n-1}\left(\frac{x-y}{\lambda}\right) = \frac{(x/\lambda)^n}{n!} e^{-x/\lambda}$$

We have thus derived another discrete distribution function, called the *Poisson distribution*, the set of probabilities  $P_n(x/\lambda)$ , that  $n$  dots, randomly placed along a line, occur in an interval of length  $x$  of this line. More generally, a sequence of events has a Poisson distribution if the outcome of an event is a positive integer and if the probability that the outcome is the integer  $n$  is

$$P_n(E) = (E^n/n!) e^{-E}; \quad \sum_{n=0}^{\infty} P_n(E) = 1 \quad (11-15)$$

If  $E < 1$ ,  $P_0(E)$  is larger than any other  $P_n$ ; if  $E > 1$ ,  $P_n$  is maximum for  $n \approx E$ , tending toward zero for  $n$  larger or smaller than this. Quantity  $E$  is the expected value of  $n$ , for

$$\langle n \rangle = \sum_{n=0}^{\infty} n P_n(E) = E \quad \text{and} \quad (\sigma_n)^2 = \sum_{n=0}^{\infty} (n - E)^2 P_n(E) = E \quad (11-16)$$

Arrivals of customers to a store during some interval of time  $T$  or the emission of particles from a radioactive source in a given interval of time—both of these have Poisson distributions. We shall encounter distributions like these later on.

### The Normal Distribution

In the limit of large values of  $N$  for the binomial distribution or of large values of  $E$  for the Poisson distribution, the variable  $n$  can be considered to be proportional to a continuous variable  $x$ , and the probability  $P_n$  approaches, in both cases, the same probability density  $F(x)$ . This function  $F(x)$  has a single maximum at  $x = \langle x \rangle$ , the expected value of  $x$  (which we shall write as  $X$ , for the time being, to save space). It drops off symmetrically on both sides of this maximum, the width of the peak being proportional to the standard deviation  $\sigma$ , as shown in Fig. 11-1. Function  $F$  has the simplest form compatible with these requirements, plus the additional requirement that the integrals representing expected values converge,

$$F(x - X) = \frac{1}{\sigma \sqrt{2\pi}} e^{-(x-X)^2/2\sigma^2}; \quad \int_{-\infty}^{\infty} F(x - X) dx = 1$$

$$\langle x \rangle = \int_{-\infty}^{\infty} x F(x - X) dx = X; \quad \langle (x - X)^2 \rangle = \sigma^2 \quad (11-17)$$

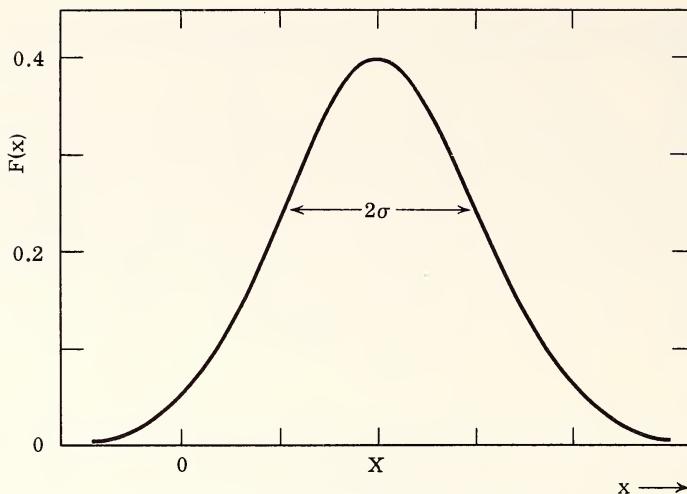


FIG. 11-1. The normal distribution

This is known as the *normal*, or Gaussian, *distribution*. It is typical of the behavior of a system subject to a large number of small, independent random effects. As an example, we might take the limiting case of a symmetric random walk ( $p = 1/2$ ), where the number  $N$  of steps is large but the size of each step is small, say  $\delta = \sigma/\sqrt{N}$ . Then, if a particular sequence of steps turns out to have had  $n$  positive steps and  $N - n$  negative ones, the net displacement would be  $x = \sigma(2n - N)/\sqrt{N}$  and the probability of such a displacement would be

$$P_n(N) = \frac{N!(1/2)^N}{[(1/2)N + (x/2\sigma)\sqrt{N}]![(1/2)N - (x/2\sigma)\sqrt{N}]!}$$

since  $n = (1/2)N + (x/2\sigma)\sqrt{N}$ .

When  $N$  is very large,  $x$  tends to become practically a continuous variable and in the limit  $P_n(N) dn \rightarrow F dx$ , where  $F$  is the probability density

$$\begin{aligned} F &= \lim_{N \rightarrow \infty} \left[ \frac{\sqrt{N}}{2\sigma} P_n(N) \right] \\ &= \lim_{N \rightarrow \infty} \left[ \frac{(\sqrt{N}/2\sigma)N!(1/2)^N}{[(1/2)N + (x/2\sigma)\sqrt{N}]![(1/2)N - (x/2\sigma)\sqrt{N}]!} \right] \end{aligned}$$

To evaluate this for large  $N$  we use the asymptotic formula for the factorial function,

$$n! \rightarrow \sqrt{2\pi n} n^n e^{-n}; \quad n > 10 \quad (11-18)$$

which is called *Stirling's formula*. Using it for each of the three factorials and rearranging factors, we obtain

$$\begin{aligned}
 F &= \lim \left[ \frac{1}{\sigma\sqrt{2\pi}} \frac{1}{\sqrt{1 - (x^2/\sigma^2N)}} \left(1 + \frac{x}{\sigma\sqrt{N}}\right)^{-(1/2)N - (x/2\sigma)\sqrt{N}} \right. \\
 &\quad \times \left. \left(1 - \frac{x}{\sigma\sqrt{N}}\right)^{-(1/2)N + (x/2\sigma)\sqrt{N}} \right] \\
 &= \lim \left[ \frac{1}{\sigma\sqrt{2\pi}} \left(1 - \frac{x^2}{\sigma^2N}\right)^{-(1/2)N - 1/2} \left(1 - \frac{x}{\sigma\sqrt{N}}\right)^{x\sqrt{N}/2\sigma} \right. \\
 &\quad \times \left. \left(1 + \frac{x}{\sigma\sqrt{N}}\right)^{-x\sqrt{N}/2\sigma} \right]
 \end{aligned}$$

By using the limiting definition of the exponential,

$$e^z = \lim_{n \rightarrow \infty} \left(1 + \frac{z}{n}\right)^n \quad (11-19)$$

this expression reduces to that of Eq. (11-17), with  $X = 0$ . The terminal point of a random walk with a large number of small steps is distributed normally, as is any other effect which is the result of a large number of independent, random components. A proof that the limiting form of the Poisson distribution also is normal constitutes one of the problems. The distribution of random errors in a series of measurements is usually normal, as is the distribution of shots at a target. In fact the normal distribution is well-nigh synonymous with the idea of randomness.

# 12 Velocity Distributions

Probability distributions are the connecting link between atomic characteristics and thermodynamic processes. We mentioned in Chapter 1 that each thermodynamic state of a system corresponded to any of a large number of microstates, macroscopically indistinguishable but microscopically different configurations of the system's atoms. If we had an assembly of thermodynamically equivalent systems, any one of the systems may be in any one of this large set of microstates; indeed each one of the systems will pass continuously from one microstate to another of the set. All we can specify is the *probability*  $f_i$  of finding the system (any one of them) in microstate  $i$ . In fact a specification of the distribution function  $f_i$ , specifying the value of  $f_i$  for each microstate possible to the system in the given macrostate, will serve to specify the thermodynamic state of each of the systems of the assembly.

## Momentum Distribution for a Gas

In the case of a perfect gas of  $N$  point atoms, each atom is equally likely to be anywhere within the volume  $V$  occupied by the gas, but the distribution-in-velocity of the atoms is less uniformly spread out. What is needed is a probability density function, prescribing the probability that an atom, chosen at random from those present, should have a specified momentum, both in magnitude and direction. We can visualize this by imagining a three-dimensional *momentum space*, as shown in Fig. 12-1, wherein the momentum  $\mathbf{p}$  of any atom can be given either in terms of its rectangular components  $p_x, p_y, p_z$  or else in terms of its magnitude  $p$  and its spherical direction angles  $\alpha$  and  $\beta$ . The probability density  $f$  is then a function of  $p_x, p_y, p_z$  or of  $p, \alpha, \beta$  (or simply of the vector  $\mathbf{p}$ ) such that  $f(\mathbf{p}) dV_p$  is the probability that an atom of gas will turn out to have a momentum vector  $\mathbf{p}$  whose head lies within the volume element  $dV_p$  in momentum space (i.e., whose  $x$  component is between  $p_x$  and  $p_x + dp_x$ ,  $y$  component

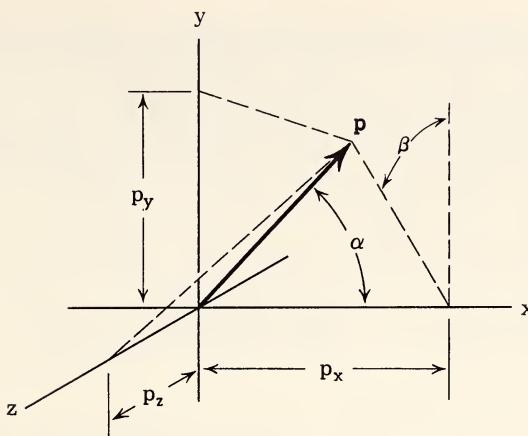


Fig. 12-1. Coordinates in momentum space.

is between  $p_y$  and  $p_y + dp_y$ , and  $z$  component is between  $p_z$  and  $p_z + dp_z$ , where  $dV_p = dp_x dp_y dp_z = p^2 dp \sin \alpha d\alpha d\beta$ .

We do not assume that any given atom keeps the same momentum forever; indeed its momentum changes suddenly, from time to time, as it collides with another atom or with the container walls. But we do assume, for the present, that these collisions are rather rare events and that if we should observe a particular atom the chances are preponderantly in favor of finding it between collisions, moving with a constant momentum, and that the probability that it has momentum  $p$  is given by the distribution function  $f(p)$  which has just been defined.

If the state of the gas is an equilibrium state we would expect that  $f$  would be independent of time; if the state is not an equilibrium one,  $f$  may depend on time as well as on  $p$ . By the basic definition of a probability density, we must have that

$$\begin{aligned} \iiint f(p) dV_p &= \int_{-\infty}^{\infty} dp_x \int_{-\infty}^{\infty} dp_y \int_{-\infty}^{\infty} dp_z f(p) \\ &= \int_0^{2\pi} d\beta \int_0^{\pi} \sin \alpha d\alpha \int_0^{\infty} f(p) p^2 dp = 1 \end{aligned} \quad (12-1)$$

since it is certain that a given atom must have some value of momentum. The distribution function will enable us to calculate all the various average values characteristic of the particular thermodynamic state specified by our choice of  $f(p)$ . For example, the mean kinetic energy of the gas molecule (between collisions, of course) is

$$\langle K.E. \rangle_{\text{tran}} = \int_0^{2\pi} d\beta \int_0^{\pi} \sin \alpha d\alpha \int_0^{\infty} f(p) (p^2/2m) p^2 dp \quad (12-2)$$

and the total energy of a gas of point atoms would be  $N < \text{K.E.} >$ , where  $N$  is the total number of atoms in the system [see Eq. (2-1)].

If the gas is moving as a whole, there will be a *drift velocity*  $V$  superimposed on the random motion, so that  $f(\mathbf{p})$  is larger in one direction of  $\mathbf{p}$  than in the opposite direction, more atoms going in the positive  $x$  direction (for example) than in the negative  $x$  direction. In this case the components of the drift velocity are

$$\begin{aligned} V_x &= \langle p_x/m \rangle = \int_{-\infty}^{\infty} (p_x/m) dp_x \int_{-\infty}^{\infty} dp_y \int_{-\infty}^{\infty} dp_z f(\mathbf{p}) \\ &= \int_0^{2\pi} d\beta \int_0^{\pi} \sin \alpha d\alpha \int_0^{\infty} (p \cos \alpha / m) f(\mathbf{p}) p^2 dp \end{aligned} \quad (12-3)$$

and similarly for  $V_y$  and  $V_z$ .

If the gas is in equilibrium and its container is at rest the drift velocity must of course be zero. In fact at equilibrium it should be just as likely to find an atom moving in one direction as in another. In other words, for a gas at equilibrium in a container at rest we should expect to find the distribution function independent of the direction angles  $\alpha$  and  $\beta$  and dependent only on the magnitude  $p$  of the momentum. In this case  $V = 0$  and

$$4\pi \int_0^{\infty} f(p) p^2 dp = 1; \quad \langle \text{K.E.} \rangle = \frac{2\pi}{m} \int_0^{\infty} f(p) p^4 dp \quad (12-4)$$

### The Maxwell Distribution

We now proceed to obtain, by a rather heuristic argument, the momentum distribution of a gas of point atoms in equilibrium at temperature  $T$ . A more "fundamental" derivation will be given later; at present understandability is more important than rigor. We have already seen that at equilibrium the distribution function should be a function of the magnitude  $p$  of the momentum, independent of the angles  $\alpha$  and  $\beta$ . One additional fact can be brought to bear: Eq. (11-3) states that, if the magnitudes of the three components of the momentum are distributed independently, then  $f(p)$  should equal the product of the probability densities of each component separately,  $f(p) = F(p_x) \cdot F(p_y) \cdot F(p_z)$ , each of the factors being similar functions of the three components.

Moreover, since the atomic motions are entirely at random, it would seem reasonable that the function  $F$  should have the form of the normal distribution, Eq. (11-17), which represents the effects of randomness. Thus we would expect that the equilibrium momentum distribution would be

$$f_0(p) = F(p_x)F(p_y)F(p_z) = \frac{1}{(2\pi\sigma^2)^{3/2}} \exp \frac{-p_x^2 - p_y^2 - p_z^2}{2\sigma^2} \quad (12-5)$$

where  $\sigma$  is the standard deviation of either of the momentum components from its zero mean value. This result strengthens our impres-

sion that we are on the right track, for the sum  $p_x^2 + p_y^2 + p_z^2 = p^2$  is independent of the angles  $\alpha$  and  $\beta$ , and thus  $f$  is a function of the magnitude of  $p$  and independent of  $\alpha$  and  $\beta$ . In order to have  $f(p)$  be, at the same time, the product of functions  $F$  of the individual components and also a function of  $p$  alone,  $f$  must have the exponential form of Eq. (12-5) (or, at least, this is the simplest function which does so).

To find the value of the variance  $\sigma^2$  in terms of the temperature of the gas, we have recourse to the results of Chapters 2 and 3, in particular of Eq. (3-2), relating the mean kinetic energy of translational motion of the atoms in a perfect gas with the temperature  $T$ . To compute mean values for the normal distribution we write down the following integral formulas:

$$\begin{aligned} \int_0^\infty e^{-u^2/a} du &= \frac{1}{2} \sqrt{\pi a} \\ \int_0^\infty e^{-u^2/a} u^{2n} du &= \frac{1}{2} \sqrt{\pi a} \quad \frac{1}{2} \cdot \frac{3}{2} \cdot \frac{5}{2} \cdots \frac{2n-1}{2} a^n \quad (12-6) \\ \int_0^\infty e^{-u^2/a} u^{2n+1} du &= \frac{1}{2} n! a^{n+1} \end{aligned}$$

Therefore the mean value of the atomic kinetic energy is

$$\begin{aligned} \langle K.E. \rangle &= 4\pi \int_0^\infty (p^2/2m)f(p)p^2 dp = \frac{1}{m\sigma^3\sqrt{2\pi}} \int_0^\infty p^4 e^{-p^2/2\sigma^2} dp \\ &= \frac{3}{2} (\sigma^2/m) \end{aligned}$$

which must equal  $(3/2)kT$ , according to Eq. (3-2). Therefore the variance  $\sigma^2$  is equal to  $mkT$ , where  $m$  is the atomic mass,  $k$  is Boltzmann's constant, and  $T$  is the thermodynamic temperature of the gas.

Thus a rather heuristic argument has led us to the following momentum distribution for the translational motion of atoms in a perfect gas at temperature  $T$ ,

$$f(p) = (2\pi mkT)^{-3/2} e^{-p^2/2mkT} \quad (12-7)$$

which is called the *Maxwell distribution*. It is often expressed in terms of velocity  $v = p/m$  instead of momentum. Experimentally we find that it corresponds closely to the velocity distribution of molecules in actual gases. The distribution is a simple one, being isotropic, with a maximum at  $p = 0$  and going to zero at  $p \rightarrow \infty$ . The integral giving the fraction of particles that have speeds larger than  $v$  is

$$4\pi \int_{mv}^{\infty} f(p)p^2 dp = \frac{2}{\sqrt{\pi}} e^{-mv^2/2kT} \int_0^{\infty} e^{-u} \left[ \frac{mv^2}{2kT} + u \right]^{1/2} du$$

$$\rightarrow 1.13 \sqrt{mv^2/2kT} e^{-mv^2/2kT}$$

when  $mv^2 \gg 2kT$

Thus about half of the atoms have speeds greater than  $\sqrt{2kT/m}$ , about  $1/25$  of them have speeds greater than  $2\sqrt{2kT/m}$ , and only one atom in about 2400 has a speed greater than  $3\sqrt{2kT/m}$ .

### Mean Values

The mean velocity of the atoms is, of course, zero, since the momentum distribution is symmetric. The mean speed and the mean-squared speed are

$$\langle v \rangle = \langle p/m \rangle = \frac{4\pi}{m} \int_0^{\infty} p^3 e^{-p^2/2mkT} \frac{dp}{(2\pi mkT)^{3/2}} = \sqrt{\frac{8kT}{\pi m}}$$

$$\langle v^2 \rangle = \langle p^2/m^2 \rangle = \frac{3}{2} (kT/m) \quad (12-8)$$

We note that the mean of the square of the speed is not exactly equal to the square of the mean speed ( $8/\pi$  is not exactly equal to  $3/2$ , although the difference is not large). The mean molecular kinetic energy of translation is proportional to  $T$ ; the mean molecular speed is proportional to  $\sqrt{T}$ .

If the gas is a mixture of two kinds of molecules, one with mass  $m_1$ , the other with mass  $m_2$ , then each species of molecule will have its own distribution in velocity, the one with  $m_1$ , instead of  $m$  in the expression of Eq. (12-7), the other with  $m_2$  instead of  $m$ . This is equivalent to saying that the mean kinetic energy of translational motion of each species is  $(3/2)kT$ , no matter what the molecular weight of each is, as long as the two kinds are in equilibrium at temperature  $T$ . In fact if a dust particle of mass  $M$  is floating in the gas, being in equilibrium with the molecules of the gas, it will be in continuous, irregular motion (called *Brownian motion*), which is equivalent to the thermal motion of the molecules, so its mean kinetic energy of translation will also be  $(3/2)kT$ . Its mean-square speed, of course, will be less than the mean-square speed of a gas molecule, by a factor equal to the square root of the ratio of the mass of the molecule to the mass of the dust particle.

Finally, we should check to make sure that a gas with molecules having a Maxwell distribution of momentum will have a pressure corresponding to the perfect gas law of Eq. (3-1). Of those molecules which have an  $x$  component of momentum equal to  $p_x$ ,  $(N/V) dA (p_x/m)$  of them will strike per second on an area  $dA$ , perpendicular to the  $x$  axis,  $N/V$  being the number per unit volume and  $(p_x/m)dA$  being the volume of the prism, in front of  $dA$ , which contains all the molecules

that will strike  $dA$  in a second. Each such molecule, striking  $dA$ , would impart a momentum  $2p_x$  if  $dA$  were a part of the container wall, so that the average momentum given to  $dA$  per second, which is equal to the pressure times  $dA$ , is (see Fig. 2-1)

$$P dA = (N/V) dA \int_{-\infty}^{\infty} dp_z \int_{-\infty}^{\infty} dp_y \int_0^{\infty} (2p_x^2/m) f(p) dp_x$$

or

$$\begin{aligned} P &= \frac{N/V}{(2\pi mkT)^{3/2}} \int_0^{\infty} e^{-p_x^2/2mkT} (2p_x^2/m) dp_x \int_{-\infty}^{\infty} \int_0^{\infty} e^{-(p_y^2 + p_z^2)/2mkT} \\ &\quad \times dp_y dp_z \\ &= \frac{N}{V} \sqrt{\frac{2}{\pi m^3 k T}} \int_0^{\infty} p_x^2 e^{-p_x^2/2mkT} dp_x = NkT/V = nRT/V \end{aligned} \quad (12-9)$$

where integration is only for positive values of  $p_x$  since only those with positive values of  $p_x$  are going to hit the area  $dA$  in the next second; the ones with negative values have already hit. Thus a gas with a Maxwellian distribution of momentum obeys the perfect gas law.

### Collisions between Gas Molecules

Most of the time a gas molecule is moving freely, unaffected by the presence of other molecules of the gas. Occasionally, of course, two molecules collide, bouncing off with changed velocities. Roughly speaking, if two molecules come within a certain distance  $R$  of each other their relative motion is affected and we say they have collided; if their centers are farther apart than  $R$  they are not affected. To each molecule, all other molecules behave like targets, each of area  $\sigma_c = \pi R^2$ , perpendicular to the path of the molecule in question. If the path of this molecule's center of mass happens to intersect one of these targets, a collision has occurred and the path changes direction. Since there are  $N/V$  molecules in a unit volume, then in a disk-like region, unit area wide and  $dx$  thick, there are  $(N/V) dx$  molecules. Therefore, the fraction of the disk obstructed by targets is  $(N\sigma_c/V) dx$  and consequently the chance of the molecule in question having a collision while moving a distance  $dx$  is  $(N\sigma_c/V) dx$ . Target area  $\sigma_c$  is called the *collision cross section* of the molecules.

Thus a collision comes at random to a molecule as it moves around; the density  $(1/\lambda)$  of their occurrence along the path of its motion is  $N\sigma_c/V$ . Reference back to the discussion before and after Eqs. (11-11) and (11-12) indicates that if the chance of encountering a "dot" (i.e., a collision) is  $dx/\lambda$ ,  $\lambda$  is the mean distance between col-

lisions (or dots) and the probability that the molecule travels a distance  $x$  without colliding and then has its next collision in the next  $dx$  of travel is

$$f(x) dx = (dx/\lambda) e^{-x/\lambda} \quad \text{where } \lambda = (V/\sigma_c N) \quad (12-10)$$

The mean distance between collisions  $\lambda$  is called the *mean free path* of the molecule. We see that it is inversely proportional to the density ( $N/V$ ) of molecules and also inversely proportional to the molecular cross section  $\sigma_c$ . This mean free path is usually considerably longer than the mean distance between molecules in a gas. For example,  $\sigma_c$  for  $O_2$  molecules is roughly  $4 \times 10^{-19} \text{ m}^2$  and  $N/V$  at standard pressure and temperature ( $0^\circ \text{C}$  and  $1 \text{ atm}$ ) is approximately  $2.5 \times 10^{25}$  molecules per  $\text{m}^3$ . The mean distance between molecules is then the reciprocal of the cube root of this, or approximately  $3.5 \times 10^{-9} \text{ m}$ , whereas the mean free path is  $\lambda = V/N\sigma_c \simeq 10^{-7} \text{ m}$ , roughly 30 times larger. The reason, of course, is that the collision radius  $R$  is about  $4 \times 10^{-10} \text{ m}$ , for  $O_2$ , about one-tenth of the mean distance between molecules; thus only about 1/1000 of the volume is "occupied" by molecules at standard conditions. The difference is even more marked at low pressures. At  $10^{-7} \text{ atmosphere}$   $\lambda$  is 1 meter in length, but there are still  $2.5 \times 10^{18}$  molecules per  $\text{m}^3$  at this pressure, hence the mean distance between molecules is roughly  $0.7 \times 10^{-6} \text{ m}$ . Neither  $\lambda$  nor  $\sigma_c$  is dependent on the velocity distribution of the molecules.

We can also talk about a mean time  $\tau$  between collisions, although this quantity does depend on the molecular velocity distribution. For a molecule having a speed  $p/m$  it would take, on the average,  $m\lambda/p$  seconds for it to go from one collision to the next. Therefore the mean free time  $\tau$  for a gas with a Maxwellian distribution of momentum is

$$\tau = \langle m\lambda/p \rangle = 4\pi m \lambda \int_0^\infty p e^{-p^2/2mkT} \frac{dp}{(2\pi mkT)^{3/2}} = \lambda \sqrt{\frac{2m}{\pi kT}} \quad (12-11)$$

which is not exactly equal to either  $\lambda/\langle v \rangle$  or  $\lambda/\sqrt{\langle v^2 \rangle}$  but it is not very different from either. The mean free time decreases with increasing temperature because an increase in temperature increases the mean molecular speed, whereas it does not change  $\lambda$ . Since the mean speed of an oxygen molecule at standard conditions is about  $400 \text{ m/sec}$ , the mean free time for these conditions is about  $3 \times 10^{-10} \text{ sec}$ ; at  $0^\circ \text{C}$  and  $10^{-7} \text{ atm}$  it is about  $(1/300) \text{ sec}$ .

# 13

## The Maxwell-Boltzmann Distributions

When the mean density of particles in a gas is not uniform throughout the gas, or when electric or gravitational forces act on the molecules, then the distribution function for the molecules depends on position as well as momentum, and may also depend on time. In such cases we must talk about the probability  $f(\mathbf{r}, \mathbf{p}, t) dV_r dV_p$  that a molecule is in the position volume element  $dV_r = dx dy dz$  at the point  $x, y, z$  (denoted by the vector  $\mathbf{r}$ ) within the container, and has a momentum in the momentum volume element  $dV_p = dp_x dp_y dp_z$  at  $p_x, p_y, p_z$  (denoted by the vector  $\mathbf{p}$ ) at time  $t$ . Because  $f$  is a probability density we must have  $\int dV_r \int dV_p f(\mathbf{r}, \mathbf{p}, t) = 1$ , where the integration over  $dV_r$  is over the interior of the container enclosing the gas and where the integration over  $dV_p$  is usually over all magnitudes and directions of  $\mathbf{p}$ .

### Phase Space

To determine the dependence of  $f$  on  $\mathbf{r}$ ,  $\mathbf{p}$ , and  $t$  we shall work out a differential equation which it must satisfy. The equation simply takes into account the interrelation between the force on a particle, its momentum, and its position in space. In addition to the time variable,  $f$  is a function of six coordinates, the three position coordinates and the three momentum coordinates. A point in this six-dimensional *phase space* represents both the position and momentum of the particle. As the particle moves about in phase space, its momentum coordinates change in accordance with the force on the particle and its position coordinates change in accordance with the momentum it has. Each molecule of the gas has its representative point in phase space; all the points move about as a swarm, their density at any time and any point in phase space being proportional to  $f(\mathbf{r}, \mathbf{p}, t)$ .

The point which is at  $(x, y, z, p_x, p_y, p_z)$  in phase space has a six-dimensional "velocity" which has components  $(\dot{x}, \dot{y}, \dot{z}, \dot{p}_x, \dot{p}_y, \dot{p}_z)$ , where the dot represents the time differential. We note the fact that the

“coordinates” are related to the “velocity” components in a rather crosswise manner, for the  $x, y, z$  part of the velocity,  $\dot{\mathbf{r}} = \mathbf{p}/m$ , is proportional to the momentum part  $\mathbf{p}$  of the phase-space coordinates, independent of  $\mathbf{r}$ ; and the momentum part of the velocity,  $\dot{\mathbf{p}} = \mathbf{F}(\mathbf{r})$ , is the particle’s rate of change of momentum, which is equal to the force  $\mathbf{F}$  on the particle, which is a function of  $\mathbf{r}$ . Thus two points in phase space which have the same space components  $x, y, z$  but different momentum coordinates  $p_x, p_y, p_z$  have “velocities” with equal momentum components  $\dot{p}_x, \dot{p}_y, \dot{p}_z$  [because  $\mathbf{F}(\mathbf{r})$  is the same for both points] but different space components  $\dot{x}, \dot{y}, \dot{z}$  (since the  $p$ ’s differ for the two points). Vice versa, two points which have the same momentum coordinates but different space coordinates will have six-dimensional velocities with the same space components  $\dot{x}, \dot{y}, \dot{z}$  but different momentum components  $\dot{p}_x, \dot{p}_y, \dot{p}_z$ .

At the instant of collision, the two points in phase space, representing the two colliding molecules, suddenly change their momentum coordinates. In other words, those two points in phase space disappear from their original positions and reappear with a new pair of momentum coordinates. Any text in hydrodynamics will show that the expression  $(\partial\rho/\partial t) + \operatorname{div} \rho\mathbf{v}$ , where  $\rho$  is the fluid density and  $\mathbf{v}$  its velocity, is equal to the net creation of fluid at the point  $x, y, z$ . Extension to phase space indicates that

$$\begin{aligned} \frac{\partial f}{\partial t} + \frac{\partial}{\partial x} (\dot{x}f) + \frac{\partial}{\partial y} (\dot{y}f) + \frac{\partial}{\partial z} (\dot{z}f) + \frac{\partial}{\partial p_x} (\dot{p}_x f) + \frac{\partial}{\partial p_y} (\dot{p}_y f) + \frac{\partial}{\partial p_z} (\dot{p}_z f) \\ = \frac{\partial f}{\partial t} + \operatorname{div}_r (\dot{\mathbf{r}}f) + \operatorname{div}_p (\dot{\mathbf{p}}f) \end{aligned} \quad (13-1)$$

measures the net appearance of points in a six-dimensional volume element at  $(x, y, z, p_x, p_y, p_z)$ , the difference between collision-produced appearances and disappearances per second. Thus the expression (13-1) should equal a function  $Q(\mathbf{r}, \mathbf{p}, t)$  called the *collision function*, the form of which is determined by the nature of the molecular collisions.

### The Boltzmann Equation

As stated two paragraphs ago, the “velocity” components  $\dot{x}, \dot{y}, \dot{z}$  are independent of  $\mathbf{r}$  and the components  $\dot{p}_x, \dot{p}_y, \dot{p}_z$  are independent of the  $p$ ’s (but depend on  $\mathbf{r}$ ). Therefore Eq. (13-1) becomes

$$\frac{\partial f}{\partial t} + \dot{x} \frac{\partial f}{\partial x} + \dot{y} \frac{\partial f}{\partial y} + \dot{z} \frac{\partial f}{\partial z} + \dot{p}_x \frac{\partial f}{\partial p_x} + \dot{p}_y \frac{\partial f}{\partial p_y} + \dot{p}_z \frac{\partial f}{\partial p_z} = Q$$

or

$$(\partial f / \partial t) + (\mathbf{p}/m) \cdot \nabla_{\mathbf{r}} f + \mathbf{F} \cdot \nabla_{\mathbf{p}} f = Q(\mathbf{r}, \mathbf{p}, t) \quad (13-2)$$

where  $\mathbf{F}(\mathbf{r})$  is the force on a gas molecule when it is at position  $\mathbf{r}$ . This equation, which is called the *Boltzmann equation*, can also be obtained by assuming that the swarm of points, representing the molecules, as it moves along in phase space, only changes its density because of the net difference between appearances and disappearances produced by collisions.

The density of the swarm at  $\mathbf{r}, \mathbf{p}, t$  is proportional to  $f(\mathbf{r}, \mathbf{p}, t)$ . A time  $dt$  later this swarm is at  $\mathbf{r} + (\mathbf{p}/m) dt$ ,  $\mathbf{p} + \mathbf{F} dt$ ,  $t + dt$ , so we should have

$$f\left(\mathbf{r} + \frac{\mathbf{p}}{m} dt, \mathbf{p} + \mathbf{F} dt, t + dt\right) - f(\mathbf{r}, \mathbf{p}, t) = Q dt$$

the difference being the net gain of points in time  $dt$ . Expansion of the first  $f$  in a Taylor's series, subtraction and division by  $dt$  results in Eq. (13-2) again.

For a gas in equilibrium, under the influence of no external forces, its molecules have a Maxwell distribution in momentum,  $f$  is independent of  $t$  and of  $\mathbf{r}$ . Since  $f$  must satisfy Eq. (13-2), we see that for this case  $Q$  must be zero for all values of  $\mathbf{r}$  and  $\mathbf{p}$ . In other words, at equilibrium, for every pair of molecules whose points are in a given region of phase space and which collide, thus disappearing from the region, there is another pair of molecules with momentum just after collision such that they appear in the region; for every molecule which loses a given momentum by collision, there is somewhere another molecule which gains this momentum by collision.

Now suppose the gas is under the influence of a conservative force, representable by a potential energy  $\phi(\mathbf{r})$ , such that  $\mathbf{F} = -\nabla_{\mathbf{r}} \phi$ . If the gas is in equilibrium, we would expect that the balance between loss and gain of points still held and that  $Q$  is zero everywhere. However now the density of the gas may differ in different parts of the container, so  $f$  may be a function of  $\mathbf{r}$  as well as  $\mathbf{p}$ . At equilibrium, though, the dependence of  $f$  on  $\mathbf{p}$  should still be Maxwellian, since the temperature is still uniform throughout the gas. In other words the distribution function should have the form  $f(\mathbf{r}, \mathbf{p}, t) = f_r(\mathbf{r}) \cdot f_p(\mathbf{p})$ , where factor  $f_p$  will have the Maxwellian form of Eq. (12-7).

Therefore for a gas in equilibrium at temperature  $T$  in the presence of a potential field  $\phi$ , the Boltzmann equation will be

$$(\mathbf{p}/m) \cdot \nabla_{\mathbf{r}} f - (\nabla_{\mathbf{p}} \phi) \cdot (\nabla_{\mathbf{p}} f) = 0$$

Because  $f = f_r f_p$  and since  $f_p$  is given by Eq. (12-7) we have  $\nabla_{\mathbf{p}} f = -(\mathbf{p}/m k T) f_r f_p$ , so the equation further simplifies,

$$(p/m) \cdot [\text{grad}_r f_r + (f_r/kT) \text{grad}_r \phi] f_p = 0$$

or

$$\text{grad}_r f_r = - (f_r/kT) \text{grad}_r \phi \quad \text{or} \quad f_r(r) = B e^{-\phi/kT} \quad (13-3)$$

Thus the distribution function for this case is

$$f(r, p) = \frac{B}{(2\pi mkT)^{1/2}} \exp \left[ -\frac{1}{kT} \left( \frac{p^2}{2m} + \phi \right) \right] \quad (13-4)$$

This is known as the *Maxwell-Boltzmann distribution*. Constant B must be adjusted so that the integral of f over all allowed regions of phase space is unity. The formula states that the probability of presence of a molecule at a point  $r, p$  in phase space is determined by its *total energy*  $(p^2/2m) + \phi(r) = H(r, p)$ ; the larger H is, the smaller is the chance that a molecule is present; the smaller T is, the more pronounced is this probability difference between points where H differs.

### A Simple Example

Suppose a gas at equilibrium at temperature T is confined to two interconnected vessels, one of volume  $V_1$ , at zero potential energy, the other of volume  $V_2$  at the lower potential energy  $\phi = -\gamma$ . The connecting tube between the two containers should allow free passage of the gas molecules, although its volume should be negligible compared to  $V_1$  or  $V_2$ . The potential difference between the vessels may be gravitational, the vessels being at different elevations above sea level; or it may be due to an electric potential difference, if the molecules possess electric charges.

In this simple case the factor  $f_r = B e^{-\phi/kT}$  of the Maxwell-Boltzmann distribution is simply B throughout volume  $V_1$  and  $B e^{\gamma/kT}$  throughout  $V_2$ . Since the integral of  $f_r$  over the total volume must be unity, we must have  $B = [V_1 + V_2 e^{\gamma/kT}]^{-1}$ . The number of molecules per unit volume in the upper container  $V_1$  is B times the total number N of molecules in the system; the density of molecules in the lower container  $V_2$  is  $N B e^{\gamma/kT}$ . Since the distribution in momentum is Maxwellian in both vessels, the pressure in each container is  $kT$  times the density of molecules there,

$$\text{For } V_1: P_1 = \frac{NkT}{V_1 + V_2 e^{\gamma/kT}}$$

$$\text{In } V_2: P_2 = \frac{NkT}{V_1 e^{-\gamma/kT} + V_2} \quad (13-5)$$

At temperatures high enough so that  $kT \gg \gamma$  the exponentials in the denominators of these expressions are nearly unity and the pressures in the two vessels are both roughly equal to  $kT$  times the mean density  $N/(V_1 + V_2)$  of molecules in the system. If the temperature is less than  $(\gamma/k)$ , however, the pressures in the two vessels differ appreciably, being greater in the one at lower potential,  $V_2$ . When  $kT \ll \gamma$  practically all the gas is in this lower container.

The mean energy of a point atom in the upper container is all kinetic, and thus is  $(3/2)kT$ ; the mean energy of a point atom in the lower vessel is  $(3/2)kT$  plus its potential energy there,  $-\gamma$ . Consequently, if the molecules are point atoms, having only translational kinetic energy, the total energy of the system is the sum of the number of molecules in each vessel times the respective mean energies,

$$U = NBV_1 \left( \frac{3}{2}kT \right) + NBV_2 e^{\gamma/kT} \left( \frac{3}{2}kT - \gamma \right)$$

$$= \frac{3}{2} NkT - \frac{N\gamma V_2}{V_1 e^{-\gamma/kT} + V_2} \quad (13-6)$$

which changes from  $N[(3/2)kT - \gamma]$  when  $kT \ll \gamma$  and all the gas is in  $V_2$  to  $N\left(\frac{3}{2}kT - \frac{V_2\gamma}{V_1 + V_2}\right)$  when  $kT \gg \gamma$  and the density is practically the same in both containers.

To compute the entropy and other thermodynamic potentials we must first recognize that there are three independent variables,  $T$ ,  $V_1$ , and  $V_2$ . The appropriate partials of  $S$  can be obtained by the procedures of Chapter 8,

$$\left( \frac{\partial S}{\partial T} \right)_{V_1 V_2} = \frac{C_V}{T} = \frac{1}{T} \left( \frac{\partial U}{\partial T} \right)_{V_1 V_2}; \quad \left( \frac{\partial S}{\partial V_1} \right)_{T V_2} = \left( \frac{\partial P_1}{\partial T} \right)_{V_1 V_2}; \quad \text{etc.}$$

Therefore,

$$S = \frac{3}{2} Nk \ln \left( \frac{T}{T_0} \right) + \frac{3}{2} Nk + S_0 + Nk \ln \left( \frac{V_1 + V_2 e^{\gamma/kT}}{V_0} \right)$$

$$- \frac{(N\gamma V_2/T)}{V_1 e^{-\gamma/kT} + V_2} \quad (13-7)$$

$$F = U - TS = -TS_0 - \frac{3}{2} NkT \ln \left( \frac{T}{T_0} \right) - NkT \ln \left( \frac{V_1 + V_2 e^{\gamma/kT}}{V_0} \right)$$

where  $S_0$  and  $V_0$  are constants of integration.

The heat capacity of the gas at constant  $V_1$  and  $V_2$ ,

$$C_{V_1 V_2} = \frac{3}{2} Nk + \frac{(N\gamma^2 V_1 V_2 / kT^2) e^{-\gamma/kT}}{(V_1 e^{-\gamma/kT} + V_2)^2}$$

is  $(3/2)kN$  both for  $kT \ll \gamma$  and for  $kT \gg \gamma$ , but for intermediate temperatures  $C_V$  is larger than this. At low temperatures nearly all the molecules are in the lower vessel and additional heat merely speeds up the molecules; at temperatures near  $\gamma/k$  the added heat must push more molecules into the upper vessel as well as speed them all up; at very high temperatures the density in the two containers is nearly equal and additional heat again serves merely to increase kinetic energy. We also note that  $P_1 = -(\partial F / \partial V_1)$ , and similarly for  $P_2$ , as required by Eqs. (8-8).

### A More General Distribution Function

The form of the Maxwell-Boltzmann distribution suggests some generalizations. In Eq. (13-4) the expression in the exponent is the total energy of position and of motion of the center of mass of the molecule, and  $f$  itself is the probability density of position and momentum of the center of mass. One obvious generalization is to put the total energy of the molecule in the exponent and to expect that the corresponding  $f$  is the probability density that each of the molecular coordinates and momenta have specified values. The position coordinates  $q$  need not be rectangular ones, they may be angles and radii, or other orthogonal curvilinear coordinates. We specify the nature of these coordinates in terms of their scale factors  $h$ , such that  $h_i dq_i$  represents actual displacement in the  $q_i$  direction (as  $r d\theta$  is displacement in the  $\theta$  direction). For rectangular coordinates  $h$  is unity; for curvilinear coordinates  $h$  may be a function of the  $q$ 's.

Suppose each molecule has  $\nu$  degrees of freedom; then it will need  $\nu$  coordinates  $q_1, q_2, \dots, q_\nu$  to specify its configuration and position in space. If the coordinates are mutually perpendicular and if  $h_i$  is the scale factor for coordinate  $q_i$ , the volume element for the  $q$ 's is  $dV_q = h_1 dq_1 h_2 dq_2 \cdots h_\nu dq_\nu$  and the kinetic energy of the molecule is

$$\langle K.E. \rangle = \frac{1}{2} \sum_{i=1}^{\nu} m_i h_i^2 \dot{q}_i^2, \quad \dot{q}_i = dq_i/dt \quad (13-8)$$

where  $m_i$  is the effective mass for the  $i$ -th coordinate (total mass or reduced mass or moment of inertia, as the case may be).

Following the procedures of classical mechanics we define the momentum  $p_i$ , conjugate to  $q_i$ , as

$$p_i = \frac{\partial}{\partial \dot{q}_i} (\text{K.E.}) = m_i h_i^2 \dot{q}_i$$

We now define the *Hamiltonian function* for the molecule as the total energy of the molecule, expressed in terms of the p's and q's,

$$H(p, q) = \sum_{i=1}^{\nu} \frac{1}{2m_i} (p_i/h_i)^2 + \phi(q) \quad (13-9)$$

where  $\phi$  is the potential energy of the molecule, expressed in terms of the q's. The h's may also be functions of the q's, but the only dependence of H on the p's is via the squares of each  $p_i$ , as written specifically in the sum. It can then be shown that the corresponding scale factors for the momentum coordinates, the other half of the  $2\nu$ -dimensional phase space for the molecule, are the reciprocals of the h's, so that the momentum volume element is  $(dp_1/h_1)(dp_2/h_2) \cdots (dp_\nu/h_\nu) = dV_p$ .

As an example, consider a diatomic molecule, with one atom of mass  $m_1$  at position  $x_1, y_1, z_1$  and another of mass  $m_2$  at  $x_2, y_2, z_2$ . We can use, instead of these coordinates, the three coordinates of the center of mass,  $x = [(m_1 x_1 + m_2 x_2)/(m_1 + m_2)]$  and similarly for y and z, plus the distance  $r$  between the two atoms and the spherical angles  $\theta$  and  $\varphi$  giving the direction of  $r$ . Then the total kinetic energy of the molecule, expressed in terms of the velocities, is

$$\frac{1}{2} (m_1 + m_2)(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + \frac{(1/2)m_1 m_2}{m_1 + m_2} (\dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \sin^2 \theta \dot{\varphi}^2)$$

so that the volume element  $dV_q = dx dy dz dr r d\theta r \sin \theta d\varphi$ . The momenta are

$$p_x = (m_1 + m_2)\dot{x} \quad \text{etc.}, \quad p_r = m_r \dot{r} \quad p_\theta = m_r r^2 \dot{\theta}$$

$$p_\varphi = m_r r^2 \sin^2 \theta \dot{\varphi}$$

where  $m_r$  is the reduced mass  $[m_1 m_2 / (m_1 + m_2)]$ . The kinetic energy expressed in terms of the p's is

$$\frac{1}{2(m_1 + m_2)} [p_x^2 + p_y^2 + p_z^2] + \frac{1}{2m_r} \left[ p_r^2 + \frac{p_\theta^2}{r^2} + \frac{p_\varphi^2}{r^2 \sin^2 \theta} \right]$$

and the volume element  $dV_p = dp_x dp_y dp_z dp_r (dp_\theta/r)(dp_\varphi/r \sin \theta)$ .

For a molecule with  $\nu$  degrees of freedom, the distribution function is

$$f(q, p) = \frac{1}{Z_q Z_p} e^{-H(p, q)/kT} \quad (13-10)$$

where

$$Z_q = \int \cdots \int e^{-\phi(q)/kT} dV_q$$

$$Z_p = \int \cdots \int \exp \left[ - \sum_{i=1}^{\nu} \frac{p_i^2/h_i^2}{2m_i kT} \right] dV_p = (2\pi kT)^{(1/2)\nu} \sqrt{m_1 m_2 \cdots m_{\nu}}$$

and where  $f(q, p) dV_q dV_p$  is the probability that the first molecular coordinate is between  $q_1$  and  $q_1 + dq_1$ , the second is between  $q_2$  and  $q_2 + dq_2$ , and so on, that the first momentum coordinate lies between  $p_1$  and  $p_1 + dp_1$ , and so on. Since the scale factors  $h_i$  are not functions of the  $p$ 's, they enter as simple constants in the integration over the  $p$ 's and thus the normalizing constant  $Z_p$  can be written out explicitly. We can also compute explicitly the mean total kinetic energy of the molecule, no matter what its position or orientation:

$$\begin{aligned} \langle K.E. \rangle_{\text{total}} &= \frac{1}{Z_q} \int \cdots \int e^{-\phi/kT} dV_q \frac{1}{Z_p} \int \cdots \int \sum_i \left( \frac{p_i^2}{2m_i h_i^2} \right) \\ &\quad \times \exp \left( - \sum_i \frac{p_i^2/h_i^2}{2m_i kT} \right) dV_p \\ &= \left( \frac{kT}{\pi^{(1/2)\nu}} \right) \int_{-\infty}^{\infty} \cdots \int \sum_i u_i^2 \exp \left( - \sum_i u_i^2 \right) \\ &\quad \times du_1 du_2 \cdots du_{\nu} = \frac{\nu}{2} kT \end{aligned} \quad (13-11)$$

### Mean Energy per Degree of Freedom

Therefore each degree of freedom of the molecule has a mean kinetic energy  $(1/2)kT$ , no matter whether the corresponding coordinate is an angle or a distance and no matter what the magnitude of the mass  $m_i$  happens to be. The thermal energy of motion of the molecule is equally distributed among its degrees of freedom. The mean value of the potential energy of course depends on the nature of the potential function  $\phi(q)$ , although a comparison with the kinetic-energy terms indicates that if the sole dependence of  $\phi$  on coordinate  $q_i$  is through a quadratic term  $(1/2)m_i \omega_i^2 q_i^2$  then the mean potential energy for this coordinate is also  $(1/2)kT$  (see below).

This brings us to an anomaly, the resolution of which will have to await our discussion of statistical mechanics. A diatomic molecule

has six degrees of freedom (three for the center of mass, one for the interatomic distance, and two angles for the direction of  $\mathbf{r}$ , as given above) even if we do not count the electrons as separate particles. We should therefore expect the total kinetic energy of a diatomic gas to be the number of molecules  $N$  times six times  $(1/2)kT$  and thus  $U$ , the internal energy of the gas, to be at least  $3NkT$  (it should be more, for there must be a potential energy, dependent on  $\mathbf{r}$ , to hold the molecule together, and this should add something to  $U$ ). However, measurements of heat capacity  $C_V = (\partial U / \partial T)_V$  show that just above its boiling point (about  $30^\circ K$ ) the  $U$  of  $H_2$  is more nearly  $(3/2)NkT$ , that between about  $100^\circ K$  and about  $500^\circ K$  the  $U$  of  $H_2$ ,  $O_2$ , and many other diatomic gases is roughly  $(5/2)NkT$  and that only well above  $1000^\circ K$  is the  $U$  for most diatomic molecules equal to the expected value of  $(6/2)NkT$  (unless the molecules have dissociated by then). The reasons for this discrepancy can only be explained in terms of quantum theory, as will be shown in Chapter 22.

### A Simple Crystal Model

Hitherto we have tacitly assumed that the coordinates  $q$  of the molecules, as used in Eq. (13-9), are universal coordinates, referred to the same origin for all molecules. This need not be so; we can refer the coordinates of each molecule to its own origin, provided that the form of the Hamiltonian for each molecule is the same when expressed in terms of these coordinates. For example, each atom in a crystal lattice is bound elastically to its own equilibrium position, each oscillating about its own origin. The motion of each atom will affect the motion of its neighbors, but if we neglect this coupling, we can consider each atom in the lattice to be a three-dimensional harmonic oscillator, each with the same three frequencies of oscillation (in Chapter 20 we shall consider the effects of the coupling, which we here neglect). In a cubic lattice all three directions will be equivalent, so that (without coupling) the Hamiltonian for the  $j$ -th atom is

$$H_j = \frac{1}{2m} (p_{xj}^2 + p_{yj}^2 + p_{zj}^2) + \frac{1}{2} m\omega^2 (x_j^2 + y_j^2 + z_j^2) \quad (13-12)$$

where  $x_j$  is the  $x$  displacement of the  $j$ -th atom from its own equilibrium position and  $p_{xj}$  is the corresponding  $m\dot{x}_j$ . Thus expressed,  $H$  has the same form for any atom in the lattice.

In this case we can redefine the Maxwell-Boltzmann probability density as follows:  $(1/Z_q Z_p) \exp(-H_j/kT)$  is the probability density that the  $j$ -th atom is displaced  $x_j, y_j, z_j$  away from its own equilibrium position and has momentum components equal to  $p_{xj}, p_{yj}, p_{zj}$ . The normalizing constants are [using Eqs. (12-6)]

$$Z_q = \left[ \int_{-\infty}^{\infty} e^{-m\omega^2 x^2/2kT} dx \right]^3 = \left[ \frac{2\pi kT}{m\omega^2} \right]^{3/2}; \quad Z_p = (2\pi mkT)^{3/2}$$

Therefore the probability density that *any* atom, chosen at random, has momentum  $p$  and is displaced a distance  $r$  from its position of equilibrium in the lattice is

$$f(r, p) = \left( \frac{\omega}{2\pi kT} \right)^3 \exp \left[ -\frac{p^2}{2mkT} - \frac{m\omega^2 r^2}{2kT} \right] \quad (13-13)$$

for the simplified model of a crystal lattice we have been assuming.

We can now use this result to compute the total energy  $U$ , kinetic and potential, of the crystal of  $N$  atoms, occupying a volume  $V$ , at temperature  $T$ . In the first place we note that even at  $T = 0$  the crystal has potential energy of over-all compression, when it is squeezed into a volume less than its equilibrium volume  $V_0$ . If the compressibility at absolute zero is  $\kappa$ , this additional potential energy is  $(V - V_0)^2/2\kappa V_0$ ; the potential energy increases whether the crystal is compressed or stretched. We should also notice that the natural frequencies of oscillation of the atoms also are affected by compression;  $\omega$  is a function of  $V$ .

The thermal energy of vibration of a typical atom in this crystal is given by the integral of  $H_f$  over all values of  $r$  and  $p$ . Again using Eqs. (12-6) we have

$$\begin{aligned} \int \dots \int H_f dV_p dV_q &= \frac{(3/2)}{m\sqrt{2\pi mkT}} \int_{-\infty}^{\infty} p^2 e^{-p^2/2mkT} dp \\ &\quad + \frac{3}{2} m\omega^2 \sqrt{\frac{m\omega^2}{2\pi kT}} \int_{-\infty}^{\infty} x^2 e^{-m\omega^2 x^2/2kT} dx \\ &= \frac{3}{2} kT + \frac{3}{2} kT = 3kT = 3(R/N_0)T \end{aligned} \quad (13-14)$$

showing that the mean thermal kinetic energy per point particle is  $(3/2)kT$  whether the particle is bound or free and that the mean thermal potential energy of a three-dimensional harmonic oscillator is also  $(3/2)kT$ , independent of the value of  $\omega$ .

Thus, for this simple model of a crystal, with  $N = nN_0$ ,

$$U = 3nRT + [(V - V_0)^2/2\kappa V_0]; \quad C_V = 3nR = T(\partial S/\partial T)_V \quad (13-15)$$

One trouble with this simple model is that it does not provide us with enough to enable us to compute the entropy. We can obtain  $(\partial S/\partial T)_V$  from Eq. (13-15) but we do not know the value of  $(\partial S/\partial V)_T =$

$(\partial P/\partial T)_V$  [see Eq. (8-12)], unless we assume the equation of state (9-1) and set  $(\partial P/\partial T)_V = \beta/\kappa$  from it. But an even more basic deficiency is that it predicts that  $C_V = 3nR$  for all values of  $T$ . As shown in Fig. 3-1 and discussed prior to Eq. (9-2),  $C_V$  is equal to  $3nR$  for actual crystals only at high temperatures; as  $T$  goes to zero  $C_V$  goes to zero. We shall show in Chapter 20 that this behavior is a quantum-mechanical effect, related to the anomaly in the heat capacities of diatomic molecules, mentioned in the paragraph after Eq. (13-11).

### Magnetization and Curie's Law

Another example of the use of the generalized Maxwell-Boltzmann distribution is in connection with the magnetic equation of state of a paramagnetic substance. Such a material has permanent magnetic dipoles of moment  $\mu$  connected to its constituent molecules or atoms. If a uniform magnetic field  $\mathfrak{B} = \mu_0 \mathfrak{H}$  ( $\mu_0$  is not a moment, it is the permeability of vacuum) is applied to the material, each molecule will have an additional potential-energy term  $-\mu \mathfrak{B} \cos \theta$  in its Hamiltonian, where  $\theta$  is the angle between the dipole and the direction of the field (and  $\mu$  is the dipole's moment). We note that the  $\mathfrak{B}$  and  $\mathfrak{H}$  used here are those acting on the individual molecules, which differ from the fields outside the substance.

The distribution function, giving the probability density that the molecule has a given momentum, position, and orientation, can be written

$$f = f_p f_q f_\theta; \quad f_\theta = (1/Z_\theta) e^{\mu \mathfrak{B} \cos \theta / kT}$$

where  $f_p$  is the momentum distribution, so normalized that the integral of  $f_p$  over all momenta is unity;  $f_q$  is the position distribution for all position coordinates except  $\theta$ , the orientation angle of the magnetic dipole ( $f_q$  is also normalized to unity). Thus the factor  $f_\theta$  gives the probability distribution for orientation of the magnetic dipole; each dipole is most likely to be oriented along the field ( $\theta = 0$ ) and least likely to be pointed against the field ( $\theta = \pi$ ). When  $kT \gg \mu \mathfrak{B}$  the difference is not great; the thermal motion is so pronounced that the dipoles are oriented almost at random. When  $kT \ll \mu \mathfrak{B}$  the difference is quite pronounced. Nearly all the dipoles are lined up parallel to the magnetic field; the thermal motion is not large enough to knock many of them askew.

The factor  $f_\theta$  can be normalized separately,

$$Z_\theta = \int_0^\pi e^{\mu \mathfrak{B} \cos \theta / kT} \sin \theta \, d\theta = \left( \frac{2kT}{\mu \mathfrak{B}} \right) \sinh \frac{\mu \mathfrak{B}}{kT}$$

and can be used separately to find the mean value of the component  $\mu \cos \theta$  of the dipole moment along the magnetic field. This, times the number of dipoles per unit volume, is, of course, the magnetic polarization  $\mathcal{P}$  of the material, as defined in Chapter 3. And, since the magnetization is  $\mathfrak{M} = \mu_0 V \mathcal{P}$ , we have

$$\begin{aligned}\mathfrak{M} &= \frac{\mu_0 N}{Z_\theta} \int_0^\pi (\mu \cos \theta) e^{\mu \mathcal{B} \cos \theta / kT} \sin \theta \, d\theta \quad (13-16) \\ &= N \mu \mu_0 \frac{(2/x) \cosh x - (2/x^2) \sinh x}{(2/x) \sinh x}, \quad x = \frac{\mu \mathcal{B}}{kT} \\ &= N \mu \mu_0 \left[ \coth \left( \frac{\mu \mathcal{B}}{kT} \right) - \left( \frac{kT}{\mu \mathcal{B}} \right) \right] \rightarrow \begin{cases} (N \mu_0 \mu^2 \mathcal{B} / 3kT), & kT \gg \mu \mathcal{B} \\ N \mu_0 \mu, & kT \ll \mu \mathcal{B} \end{cases}\end{aligned}$$

Thus we see that at low fields (or high temperatures) the dipoles tend only slightly to line up with the field and the magnetization  $\mathfrak{M}$  is proportional to  $\mathcal{B} = \mu_0 \mathcal{C}$ , but that at high magnetic intensities (or low temperatures) all the dipoles line up and the magnetization reaches its asymptotic value  $N \mu_0 \mu$  (i.e., it is saturated), as shown in Fig. 13-1. Since the magnetic moment of an oxygen molecule is roughly

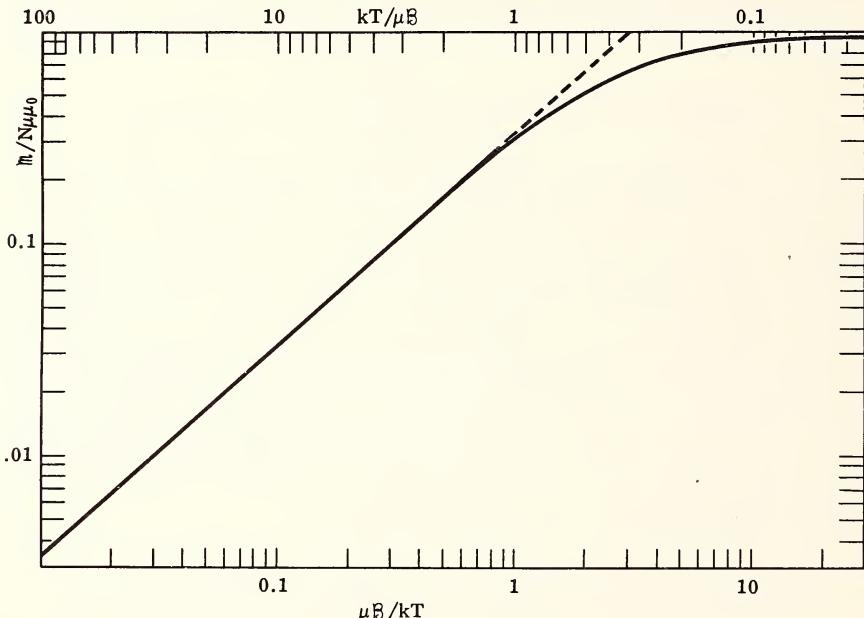


FIG. 13-1. Magnetization curve for paramagnetic substances.

$3 \times 10^{-23}$  mks units,  $\mu\mathcal{B}$  is no larger than  $3 \times 10^{-23}$  joules for  $\mathcal{B} = 1$  weber per  $\text{m}^2$  ( $= 10,000$  gauss, a quite intense field). Since  $kT = 3 \times 10^{-23}$  joules for  $T \approx 2^\circ\text{K}$ , the parameter  $x = \mu\mathcal{B}/kT$  is considerably less than unity for  $\text{O}_2$  (for example) for temperatures greater than  $30^\circ\text{K}$  and/or  $\mathcal{B}$  less than 1 weber/ $\text{m}^2$ . In such cases, where  $x \ll 1$ , the polarization  $\mathcal{P}$  is much smaller than  $\mathcal{B}$ , so that the  $\mathcal{B}$  acting on the molecule is not much different from the  $\mathcal{B}$  outside the material.

Thus for most temperatures and field strengths, for paramagnetic materials like  $\text{O}_2$ ,  $x$  is very small and Curie's law

$$\mathcal{M} = nD\mathcal{B}/T; \quad D = N_0 \mu_0^2 \mu^2 / 3k \quad (13-17)$$

is a good approximation for the magnetic equation of state. Kinetic theory has thus not only derived Curie's law [see Eq. (3-8)] and obtained a relation between the Curie constant  $D$  and the molecular characteristics of the material (such as  $\mu$  and the basic constants  $N_0$ ,  $\mu_0$ , and  $k$ ) but has also determined the limits beyond which Curie's law is no longer valid, and the equation of state which then holds.

For example, for the paramagnetic perfect gas of Chapter 7, the more accurate equation for  $T dS$  and the adiabatic formula (7-14) is

$$T dS = \left[ \frac{5}{2} nR - \frac{3nD}{\alpha^2} (x^2 \operatorname{csch}^2 x - 1) \right] dT - \frac{nRT}{P} dP + \frac{3nDT}{\alpha^2 \mathcal{B}} (x^2 \operatorname{csch}^2 x - 1) d\mathcal{B}$$

and

$$\frac{T^{5/2}}{P} \left[ \frac{T}{\alpha\mathcal{B}} \sinh \left( \frac{\alpha\mathcal{B}}{T} \right) \right]^{3D/R\alpha^2} \exp \left[ - \frac{3D\mathcal{B}}{R\alpha T} \coth \left( \frac{\alpha\mathcal{B}}{T} \right) \right] = \text{const.} \quad (13-18)$$

where  $x = \alpha\mathcal{B}/T$  and  $\alpha = \mu\mu_0/k$ . This reduces to Eq. (7-14) when  $x$  is small.

This completes our discussion of paramagnetic materials. In the case of ferromagnetic materials, where the polarization  $\mathcal{P}$  is not small compared to  $H$ , the field acting on the individual dipole is considerably modified by the polarization of the nearby dipole; in fact the dipoles may tend to line up all by themselves. But it would go too far afield to discuss permanent magnetization.

# 14

## Transport Phenomena

The Boltzmann equation (13-2) can also be used to calculate the progress of some spontaneous processes, such as the mixing of two gases by interdiffusion and the attenuation of turbulence in a gas by the action of viscosity. All these processes have to do with the transport of something, foreign molecules or electric charge or momentum, from one part of the gas to another; consequently they are called *transport phenomena*. Here, because the system is not in equilibrium, the collision term  $Q$ , which measures the net rate of entrance and exit of molecular points in phase space, is not zero. An exact calculation of the dependence of  $Q$  on  $p$  and  $r$  is quite a difficult task, requiring a detailed knowledge of molecular behavior during a collision. Exact expressions for  $Q$  have been determined for only a few, simple cases. Luckily there is a relatively simple, approximate expression for  $Q$ , which will be good enough for the calculations of this chapter.

### An Approximate Collision Function

The function  $Q$  in the Boltzmann equation represents the effects of collisions in bringing the gas into equilibrium. As we pointed out earlier, when the gas is in equilibrium  $Q$  is zero; as many molecules gain a given momentum per second by a collision as there are those that lose this momentum per second by a collision. And if the gas is close to equilibrium  $Q$  should be small. Suppose the solution of the Boltzmann equation for equilibrium conditions is  $f_0(r,p)$  and the solution for nonequilibrium conditions close to the equilibrium state is  $f(r,p)$ . What we have just been saying, in regard to the behavior of  $Q$  near equilibrium, is that it should be proportional to  $f - f_0$ , when  $f$  is nearly equal to  $f_0$ . A glance at Eq. (13-2) indicates that the proportionality constant has the dimensions of the reciprocal of time and can be written as  $1/t_c(r,p)$ , where  $t_c$  is called the *relaxation time* of the system, for reasons shortly to be apparent. We will thus assume that the Boltzmann equation, for conditions near equilibrium, can be written

$$(\partial f / \partial t) + (\mathbf{p}/m) \cdot \nabla_{\mathbf{r}} f + \mathbf{F} \cdot \nabla_{\mathbf{p}} f = [(f_0 - f)/t_c]$$

where  $f_0$  is the distribution for the nearest equilibrium state, and where  $f_0 - f$  is small compared to either  $f$  or  $f_0$ . Therefore the left-hand side of this equation is also small, and it would not produce additional first-order error to substitute  $f_0$  for  $f$  on the left-hand side. Thus an approximate equation for near-equilibrium conditions is

$$f \approx f_0 - t_c [(\partial f_0 / \partial t) + (\mathbf{p}/m) \cdot \nabla_{\mathbf{r}} f_0 + \mathbf{F} \cdot \nabla_{\mathbf{p}} f_0] \quad (14-1)$$

Collisions between gas molecules are fairly drastic interruptions of the molecule's motion. After the collision the molecule's direction of motion, and also its speed, may be drastically altered. Roughly speaking, it is as though each collision caused the participating molecule to forget its previous behavior and to start away as part of an equilibrium distribution of momentum; only later in its free time, before its next collision, does the nonequilibrium situation have a chance to reaffect its motion.

For example, we may find a gas of uniform density having initially a distribution in momentum  $f_i(\mathbf{p})$  of its molecules which differs from the Maxwell distribution (12-7); there may be more fast particles in relation to slow ones than (12-7) requires, or there may be more going in the  $x$  direction than in the  $y$ , or some other asymmetry of momentum which is still uniform in density. Such a distribution  $f_i$ , since it is not Maxwellian, is not an equilibrium distribution. However, since the lack of equilibrium is entirely in the "momentum coordinate" part of phase space, the distribution can return to equilibrium in one collision time; we would expect that at the next collision each molecule would regain its place in a Maxwell distribution, so to speak. The molecules do not all collide at once, the chance that a given molecule has not had its next collision, after a time  $t$ , is  $e^{-t/\tau}$ , where  $\tau$  is the mean free time between collisions [see Eqs. (11-11) and (12-11)]. Thus we would expect that our originally anisotropic distribution would "relax" from  $f_i$  back to  $f_0$  with an exponential dependence on time of  $e^{-t/\tau}$  (note that  $\tau$  is proportional to  $\lambda$ , the molecular mean free path).

But if  $f$  is independent of  $\mathbf{r}$  and there is no force  $\mathbf{F}$  acting, a solution of Eq. (14-1) which starts as  $f = f_i$  at  $t = 0$  is

$$f = f_0 + (f_i - f_0)e^{-t/t_c} \quad (14-2)$$

which has just the form we persuaded ourselves it should have, except that the relaxation time  $t_c$  is in the exponent, rather than the mean free time  $\tau = \langle m\lambda/p \rangle$ . We would thus expect that the relaxation time  $t_c$ , entering Eq. (14-1) would be approximately equal to  $m\lambda/p$ . Detailed calculations for the few cases which can be carried out, plus indirect experimental checks (described later in this chapter) indicate that it is not a bad approximation to set  $t_c = \langle m\lambda/p \rangle = \tau$ . This will be done in the rest of this chapter.

### Electric Conductivity in a Gas

Suppose that a certain number  $N_i$  of the molecules of a gas are ionized ( $N_i$  being small compared to the total number  $N$  of molecules) and suppose that initially the gas is at equilibrium at temperature  $T$ . At  $t = 0$  a uniform electric field  $\varepsilon$ , in the positive  $x$  direction, is turned on. The ions will then experience a force  $e\varepsilon$  in the  $x$  direction, where  $e$  is the ionic charge. Imposed on the random motion of the ions between collisions will be a "drift" in the  $x$  direction. This is not an equilibrium situation, since the drift velocity of the ions will heat up the gas. But if  $N_i/N$  is small, and if  $\varepsilon$  is small enough, the heating will be slow and we can neglect the term  $\partial f/\partial t$  in Eq. (14-1) in comparison with the other terms.

Since the ions are initially uniformly distributed in space and since the ionic drift is slow, we can assume that  $f$  is more-or-less independent of  $\mathbf{r}$ . Thus Eq. (14-1) for the ions becomes

$$f \simeq f_0 - t_c \mathbf{F} \cdot \mathbf{grad}_p f_0 = f_0 - e t_c \varepsilon (\partial f_0 / \partial p_x) \simeq [1 + e t_c \varepsilon / (m k T) p_x] f_0 \quad (14-3)$$

where  $f_0$  is the Maxwell distribution,

$$f_0 = [1 / (2 \pi m k T)^{3/2}] \exp \left[ \frac{(-p_x^2 - p_y^2 - p_z^2)}{2 m k T} \right]$$

of the neutral molecules. Function  $f$  will be a good approximation to the correct momentum distribution of the ions if the second term in the brackets of Eq. (14-3) is small compared to the first term, unity, over the range of values of  $p_x$  for which  $f_0$  has any appreciable magnitude. The term  $e t_c \varepsilon p_x / m k T$  can be written  $(e \lambda \varepsilon / k T) (p_x < 1/p>)$  if we assume that  $t_c = \tau = \langle m \lambda / p \rangle$ ,  $\lambda$  being the mean free path of the molecule [see Eq. (12-11)]. Since  $e \lambda \varepsilon$  is the energy that would be gained by the ion (in electron volts, if the ion is singly ionized) by falling through a mean free path in the direction of  $\varepsilon$ , and since  $k T$  in electron volts is  $T/7500$ , then for a gas (such as  $O_2$ ) at standard conditions, where  $\lambda \simeq 10^{-7}$  m [see the discussion following Eq. (12-10)] and  $T \simeq 300$ , the factor  $e \lambda \varepsilon / k T \simeq \varepsilon / 40,000$ ,  $\varepsilon$  being in volts per meter. Thus if  $\varepsilon$  is as large as 4000 volts per meter, the second term will not equal the first in Eq. (14-3) until  $p_x$  is 10 times the mean momentum  $\langle p \rangle$  and by this time the exponential factor of  $f_0$  will equal about  $e^{-50}$ . Thus, for a wide range of values of  $T$  and of  $\varepsilon$ , either  $f$  is vanishingly small or else the second term in brackets of Eq. (14-3) is small compared to the first.

What Eq. (14-3) indicates is that the momentum distribution of the ions, in the presence of the electric field, is slightly nonisotropic; somewhat more of them are going in the direction of the field ( $p_x$  pos-

itive) than are going in the opposite direction ( $p_x$  negative). There is a net *drift velocity* of the ions in the  $x$  direction:

$$U_x = \int_{-\infty}^{\infty} \int \int (p_x/m) f \, dV_p \simeq \int_{-\infty}^{\infty} \int \int \left( \frac{p_x}{m} + \frac{2et_c \varepsilon}{mkT} - \frac{p_x^2}{2m} \right) f_0(p) \, dV_p$$

$$\simeq \frac{et_c \varepsilon}{m} \simeq \frac{e\tau \varepsilon}{m} = \lambda e \varepsilon \left( \frac{2}{\pi mkT} \right)^{1/2} \simeq \frac{\lambda e \varepsilon}{m \langle v \rangle} = M \varepsilon \quad (14-4)$$

where we have used the fact that for a Maxwell distribution  $f_0, \langle p_x \rangle = 0$  and  $\langle p_x^2/2m \rangle = (1/2)kT$  and we have also used Eq. (12-11) for the mean free time  $\tau$ .

We see that the drift velocity  $U$  of the ion is proportional to the electric intensity, as though the ion were moving through a viscous fluid. The proportionality factor  $M \simeq et_c/m \simeq \lambda e/\langle v \rangle$  is called the *mobility* of the ion. The current density  $I = (N_i e U/V)$  (in amperes per square meter) is

$$I \simeq (N_i \lambda e^2/V) (2/\pi mkT)^{1/2} \varepsilon \quad (14-5)$$

obeying Ohm's law, with a conductivity  $N_i e M/V = N_i e^2 t_c/mV$ .

### Drift Velocity

It is interesting to see that the drift velocity, and therefore the current density, is proportional to the mean free time between collisions and is thus *inversely* proportional to the square root of the temperature  $T$ . As  $T$  increases, the *random* velocity  $\langle v \rangle$  of the ions (and neutral molecules) increases, but the *drift* velocity  $U$  of the ions decreases. One can visualize the process by imagining the flight of an ion from one collision to its next. Just after each collision the ion comes away in a random direction, with no initial preference for the direction of  $\varepsilon$ . But during its free flight the electric field acts on it, turning its motion more and more in the positive  $x$  direction (its path is a portion of a parabola) and thus adding more and more positive  $x$  component to its velocity. This accentuation of the positive  $x$  motion is completely destroyed by the next collision (on the average) and the molecule starts on a new parabolic path. If the mean free time is long, the molecule has plenty of time to add quite a bit of excess  $v_x$ ; if  $\tau$  is small, the molecule hardly has time to get acted on by the field before it collides again. Thus the higher the temperature, the greater the random velocity  $\langle v \rangle$ , the shorter the mean free time  $\tau$  and the smaller the drift velocity and current density. This, of course, checks with the measurements of gaseous conduction.

In most ionized gases, free electrons will be present as well as positive ions. The electrons will also have a drift velocity, mobility,

and current density, given by Eqs. (14-4) and (14-5), only with a negative value of charge  $e$ , a different value of  $\lambda$  and a much smaller value of mass  $m$ . Thus the drift velocity will be opposite in direction to that of the ions but, since the charge  $e$  is negative, the *current* density is in the same direction as that of the ions. Since the electronic mean free path is roughly 2 to 4 times that of the ions and since the electronic mass is several thousand times smaller, the electronic mobility is 500 to 1000 times greater than that of the positive ions and therefore most of the current in an ionized gas is carried by the electrons.

### Diffusion

Another nonequilibrium situation is one in which different kinds of molecules mix by diffusion. To make the problem simple, suppose we have a small number  $N_i$  of radioactive "tagged" molecules in a gas of  $N$  nonradioactive molecules of the same kind. Suppose, at  $t = 0$ , the distribution in space of the tagged molecules is not uniform (although the density of the mixture *is* uniform). Thus the distribution function for the tagged molecules is a function of  $\mathbf{r}$ , and we have to write our "0-th approximation" as

$$f_0 = f_r(\mathbf{r}) \cdot i_p(p); \quad f_p = \frac{e^{-p^2/2mkT}}{(2\pi mkT)^{3/2}}; \quad \int \int \int f_r dV_r = 1 \quad (14-6)$$

The distribution function  $f$  for the diffusing molecules will change with time, but we will find that the rate of diffusion is slow enough so that the term  $\partial f / \partial t$  in Eq. (14-1) is negligible compared to other terms.

In the case of diffusion there is no force  $\mathbf{F}$ , but  $f_0$  does depend on  $\mathbf{r}$ , so the approximate solution of Eq. (14-1) is

$$f \simeq f_0 - t_c(p/m) \cdot \text{grad}_r f_0 = [f_r - (t_c/m)p \cdot \text{grad}_r f_r]f_p \quad (14-7)$$

where the vector  $\text{grad}_r f_r$  points in the direction of increasing density of the tagged molecules. The anisotropy is again in the momentum distribution, but here the preponderance is opposite to  $\text{grad}_r f_r$ , there is a tendency of the tagged molecules to flow *away* from the region of highest density. The conditional probability density that a molecule, if it is at point  $\mathbf{r}$ , has a momentum  $\mathbf{p}$ , is [see Eq. (11-2)]

$$(f/f_r) \simeq [1 - (t_c/m)p \cdot \mathbf{g}]f_p; \quad \mathbf{g} = (1/f_r) \text{grad}_r f_r$$

From this we can compute the mean drift velocity of the tagged molecules which are at point  $\mathbf{r}$  (for convenience we point the  $x$  axis in the direction of  $\mathbf{g}$ ):

$$\begin{aligned}
 \mathbf{U} &\simeq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\mathbf{p}/m) [1 - (t_c/m) p_X g] f_p dV_p = -2(t_c/m) g \\
 &\times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (p_X^2/2m) f_p dV_p \simeq -(t_c kT/m) g \simeq -\lambda (2kT/\pi m)^{1/2} g \\
 &\simeq -2\lambda \langle v \rangle (1/f_r) \text{grad}_r f_r
 \end{aligned}$$

We see that in this case the drift velocity increases as  $T$  increases.

The density  $\rho_i$  of tagged molecules at  $\mathbf{r}$  is  $N_i f_r$  molecules per unit volume, so the flux  $\mathbf{J}$  of tagged particles at  $\mathbf{r}$ , the net diffusive flow caused by the uneven distribution of these particles, is

$$\mathbf{J} = N_i f_r \mathbf{U} \simeq -D \text{grad}_r \rho_i; \quad D = t_c kT/m \simeq \lambda (2kT/\pi m)^{1/2} \quad (14-8)$$

where constant  $D$  is called the *diffusion constant* of the tagged molecules. A density gradient of tagged molecules produces a net flow away from the regions of high density, the magnitude of the flow being proportional to the diffusion constant  $D$ . We note that there is a simple relationship between  $D$  and the mobility  $M$  of the same molecule when ionized and in an electric field, as given in Eq. (14-4),

$$D = (kT/e)M \quad (14-9)$$

which is more accurate than our approximation for  $t_c$ .

Thus a measurement of diffusion in a gas enables us to predict the electrical conductivity of the gas or vice versa. Equation (14-8) is the basic equation governing diffusion. By adding to it the equation of continuity, we obtain

$$\partial \rho_i / \partial t = -\text{div} \mathbf{J} \simeq D \nabla^2 \rho_i \quad (14-10)$$

which is called the *diffusion equation*.

There are a number of other transport problems, heat flow and viscosity, for example, which can be worked out by use of the Boltzmann equation. These will be given as problems.

# 15

## Fluctuations

Any system in thermal equilibrium with its surroundings undergoes fluctuations in position and velocity because of the thermal motion of its own molecules as well as of any molecules that may surround it. Kinetic theory, which enables us to compute the mean thermal kinetic and potential energy inhering in each degree of freedom of the system, makes it possible to compute the variance (i.e., the mean-square amplitude of the fluctuations) of each coordinate and momentum of the system. It is often useful to know the size of these variances, for they tell us the lower bound to the accuracy of a piece of measuring equipment and they sometimes give us a chance to measure, indirectly, the magnitude of some atomic constants, such as Avogadro's number  $N_0$ .

### Equipartition of Energy

Referring to Eqs. (13-9) and (13-10), we see that if the Hamiltonian function for a system can be separated into a sum of terms  $[(1/2m_i) \times (p_i/h_i)^2 + \phi_i(q_i)]$ , each of which is a function of just one pair of variables,  $p_i$  and  $q_i$ , then the Maxwell-Boltzmann probability density can be separated into a product of factors,  $(1/Z_i) \exp[-(1/2m_i kT)(p_i/h_i)^2 - (1/kT)\phi_i(q_i)]$ , each of which gives the distribution in momentum and position of one separate degree of freedom. Even if the potential energy, or the scale factors  $h$ , cannot be completely separated for all the degrees of freedom of the system, if the potential energy does not depend on some coordinates  $q_j$  (such as the  $x$  coordinates of the center of mass of a dust particle floating freely in the air), then all values of that coordinate are equally likely (the dust particle can be anywhere in the gas) and its momentum will be distributed according to the probability density

$$f_{pj}(p_j) = (1/Z_j) \exp \left[ -\frac{1}{2m_j kT} \left( \frac{p_j}{h_j} \right)^2 \right]; \quad Z_j = (2\pi m_j kT)^{1/2} \quad (15-1)$$

The mean thermal kinetic energy of the  $j$ -th degree of freedom is thus

$$\langle K.E. \rangle = \int_{-\infty}^{\infty} (p_j^2/2m_j h_j^2) f_{pj} (dp_j/h_j) = \frac{1}{2} kT \quad (15-2)$$

whether the coordinate is an angle or a distance or some other kind of curvilinear coordinate. Therefore the kinetic energy of thermal motion is equally apportioned, on the average, over all separable degrees of freedom of the system, an energy  $(1/2)kT$  going to each. If the potential energy is independent of  $q_j$ , then  $(1/2)kT$  is the total mean energy possessed by the  $j$ -th degree of freedom. On the average the energy of rotation of a diatomic molecule (described in terms of two angles) would be  $kT$  and the average energy of translation of its center of mass would be  $(3/2)kT$ . A light atom (helium for example) will have a higher mean speed than does a heavy atom (xenon for example) at the same temperature, in order that the mean kinetic energy of the two be equal. In fact the mean-square value of the  $j$ -th velocity, when  $q_j$  is a rectangular coordinate (i.e., when  $h_j = 1$ ), is

$$\langle \dot{q}_j^2 \rangle = \langle p_j^2/m_j^2 \rangle = kT/m_j \quad (15-3)$$

### Mean-Square Velocity

For example, the  $x$  component of the velocity of a dust particle in the air fluctuates irregularly, as the air molecules knock it about. The average value of  $\dot{x}$  is zero (if the air has no gross motion) but the mean-square value of  $\dot{x}$  is just  $kT$  divided by the mass of the particle. We note that this mean-square value is independent of the pressure or density of the air the particle is floating in, and is thus independent of the number of molecules which hit it per second. If the gas is rarefied only a few molecules hit it per second and the value of  $\dot{x}$  changes only a few times a second; if the gas is dense the collisions occur more often and the velocity *changes* more frequently per second (as shown in Fig. 15-1), but the mean-square value of the velocity is the same in both cases if the temperature is the same.

Even if the potential energy does depend on the coordinate  $q_j$ , the mean kinetic energy of the  $j$ -th degree of freedom is still  $(1/2)kT$ . If the potential energy can be separated into a term  $\phi_j(q_j)$  and another term which is independent of  $q_j$ , then the probability density that the  $j$ th coordinate has a value  $q_j$  is

$$f_{qj}(q_j) = (1/Z_{qj})e^{-\phi_j/kT}; \quad Z_{qj} = \int e^{-\phi_j/kT} h_j dq_j \quad (15-4)$$

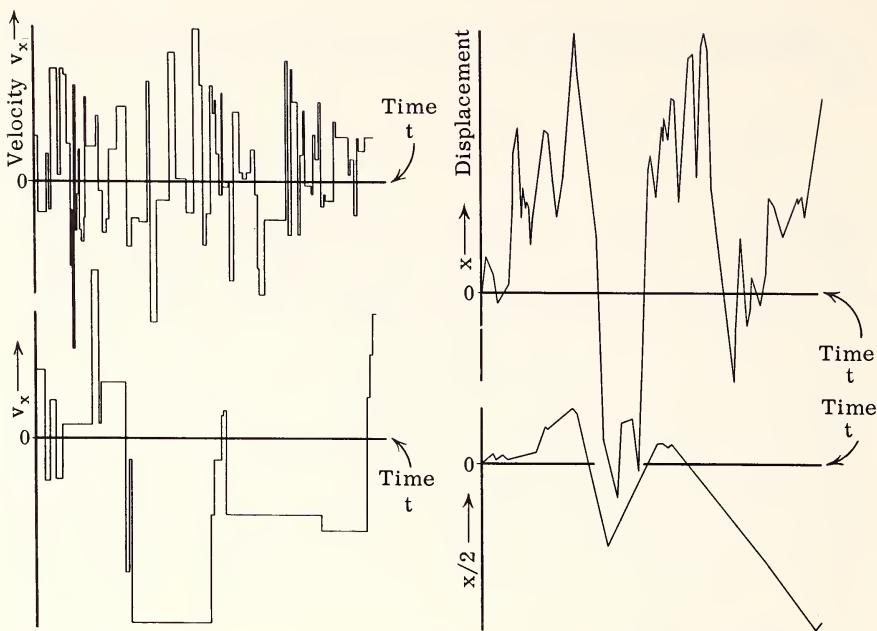


Fig. 15-1. Variation with time of  $x$  component of velocity and displacement of Brownian motion. Lower curves for mean time between collisions five times that for upper curves.

where the integration is over the allowed range of  $q_j$ . The mean value of the potential energy turns out to be a function of  $kT$ , but the nature of the function depends on how  $\phi_j$  varies with  $q_j$ .

The usual case is the one where the scale constant is unity and where the potential energy  $\phi_j = (1/2)m_j \omega_j^2 q_j^2$  has a quadratic dependence on  $q_j$ , so that in the absence of thermal motion the displacement  $q_j$  executes simple harmonic motion with frequency  $\omega_j/2\pi$ . In this case  $Z_{qj} = (2\pi kT/m_j \omega_j^2)^{1/2}$ , and the mean value of potential energy when thermal fluctuations are present is

$$\langle \phi_j \rangle = \frac{1}{2} m_j \omega_j^2 \int_{-\infty}^{\infty} q_j^2 f_{qj} dq_j = \frac{1}{2} kT \quad (15-5)$$

Thus when the potential energy per degree of freedom is a quadratic function of each  $q$ , the mean potential energy per  $q$  is  $(1/2)kT$ , independent of  $\omega_j$ , and equal to the mean kinetic energy, so that the total mean energy for this degree of freedom is  $kT$ . Also the mean-square displacement of the coordinate from its equilibrium position is

$$\langle q_j^2 \rangle = kT/m_j \omega_j^2$$

for coordinates with quadratic potentials.

As an example of a case where  $\phi_j$  is not quadratic, we recall the case of the magnetic dipoles, where  $\phi_j = -\mu \mathcal{B} \cos \theta$ , so that  $Z_{qj} = (2kT/\mu \mathcal{B}) \sinh(\mu \mathcal{B}/kT)$ . From Eq. (13-16) the mean potential energy is

$$\langle -\mu \mathcal{B} \cos \theta \rangle = kT - \mu \mathcal{B} \coth(\mu \mathcal{B}/kT) \rightarrow \begin{cases} -\frac{1}{3}(\mu^2 \mathcal{B}^2/kT), & kT \gg \mu \mathcal{B} \\ -\mu \mathcal{B} + kT, & kT \ll \mu \mathcal{B} \end{cases}$$

which is not equal to  $(1/2)kT$ .

### Fluctuations of Simple Systems

A mass  $M$ , on the end of a spring of stiffness constant  $K$ , is constantly undergoing small, forced oscillations because of thermal fluctuations of the pressure of the gas surrounding it and also because of thermal fluctuations of the spring itself. In the absence of these fluctuations the mass will describe simple harmonic motion of amplitude  $A$  and frequency  $(1/2\pi)(K/M)^{1/2}$ , so its displacement, and its mean kinetic and potential energy would be

$$x = A \cos[t(K/M)^{1/2} + \alpha]; \quad \langle \text{K.E.} \rangle = \frac{1}{4}KA^2 = \langle \text{P.E.} \rangle$$

where  $A$  and  $\alpha$  are determined by the way the mass is started into motion. In the presence of the thermal fluctuations an irregular motion is superposed on this steady-state oscillation. Even if there is no steady-state oscillation the mass will never be completely at rest but will exhibit a residual motion having total energy, potential plus kinetic, of  $kT$ , having a mean-square amplitude such that  $(1/2)KA_T^2 = kT$ , or

$$A_T^2 = 2kT/K.$$

With a mass of a few grams and a natural frequency of a few cycles per second ( $K \approx 1$ ), this mean-square amplitude is very small, of the order of  $10^{-20} \text{ m}^2$ , a root-mean-square amplitude of about  $10^{-8} \text{ cm}$ . This is usually negligible, but in some cases it is of practical importance. The human eardrum, plus the bony structure coupling it to the inner ear, acts like a mass-spring system. Even when there is no noise present the system fluctuates with thermal motion having a mean amplitude of about  $10^{-8} \text{ cm}$ . Sounds so faint that they drive the eardrum with less amplitude than this are "drowned out" by the thermal noise. In actual fact this thermal-noise motion of the eardrum sets the lower limit of audibility of sounds in the frequency range of greatest sensitivity of

the ear (1000 to 3000 cps); if the incoming noise level is less than this "threshold of audibility," we "hear" the thermal fluctuations of our eardrums rather than the outside noises.

We notice that the root-mean-square amplitude of thermal motion of a mass on a spring,  $(2kT/K)^{1/2}$ , is independent of the density of the ambient air and thus independent of the number of molecular blows impinging on the mass per second. If the density is high the motion will be quite irregular because of the large number of blows per second; if the density is low the motion will be "smoother," but the mean-square amplitude will be the same if the temperature is the same, as illustrated in Fig. 15-2. Even if the mass-spring system is in a vacuum the motion will still be present, caused by the thermal fluctuations of the atoms in the spring.

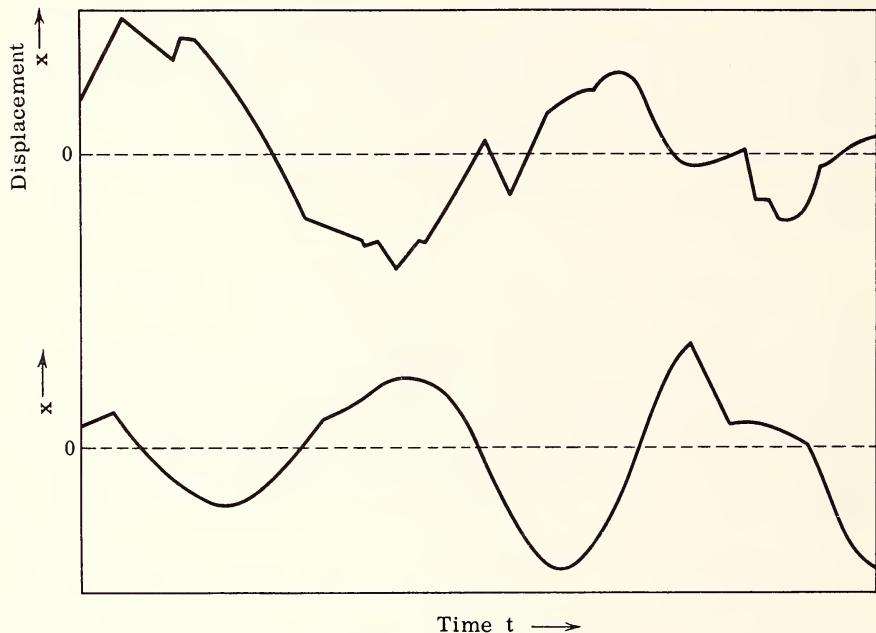


Fig. 15-2. Brownian motion of a simple oscillator for two different mean times between collisions.

The same effect is present in more-complex systems, each degree of freedom has mean kinetic energy  $(1/2)kT$  and similarly for the potential energy if it depends quadratically on the displacement, as in a simple oscillator. A string of mass  $\rho$  per unit length and length  $L$  under tension  $Q$  can oscillate in any one of its standing waves; the displacement from equilibrium and the total energy of vibration of the  $n$ -th wave is

$$y_n = A_n \sin(\pi n x / L) \cos[(\pi n t / L)(Q/\rho)^{1/2} + \alpha_n]$$

$$\langle \text{K.E.} \rangle + \langle \text{P.E.} \rangle = \frac{1}{2} Q A_n^2 (\pi n / L)^2 \int_0^L \sin^2(\pi n x / L) dx$$

$$= \frac{1}{4} (\pi^2 n^2 Q / L) A_n^2$$

When the string is at rest except for its thermal vibrations, each of the standing waves has a mean energy of  $kT$ , so the mean-square amplitude of motion  $A_n$  of the  $n$ -th wave is  $(4LkT/\pi^2 n^2 Q)$  and the mean-square amplitude of deflection of some central portion of the string is the sum of all the standing waves (because of the incoherence of the motion, we sum the squares),

$$\langle y^2 \rangle \simeq \frac{1}{4} \sum_{n=1}^{3N} \langle A_n^2 \rangle = (LkT/Q) \sum_{n=1}^{3N} (1/\pi^2 n^2) \simeq LkT/6Q$$

which is related to the result for the simple oscillator,  $12Q/L$  being equivalent to the stiffness constant  $K$  of the simple spring. If the string is part of a string galvanometer these thermal fluctuations will mask the detection of any current that deflects the string by an amount less than  $(LkT/6Q)^{1/2}$ .

### Density Fluctuations in a Gas

The thermal motion of the constituent molecules produces fluctuations of density, and thus of pressure, in a gas. We could analyze the fluctuation in terms of pressure waves in the gas, as we analyzed the motion of a string under tension in terms of standing waves. Instead of this, however, we shall work out the problem in terms of the potential energy of compression of the gas. Suppose we consider that part of the gas which, in equilibrium, would occupy a volume  $V_S$  and would contain  $N_S$  molecules. At temperature  $T$  the gas, at equilibrium, would have a pressure  $P = N_S kT / V_S$  throughout. If the portion of the gas originally occupying volume  $V_S$  were compressed into a somewhat smaller volume  $V_S - \Delta V$  ( $\Delta V \ll V_S$ ), an additional pressure  $\Delta P = [N_S kT / (V_S - \Delta V)] - P \simeq P(\Delta V / V_S)$  would be needed and an amount of work

$$\int_0^{\Delta V} \Delta P d(\Delta V) = \frac{1}{2} (P/V_S)(\Delta V)^2 = \frac{1}{2} N_S kT (\Delta V/V_S)^2$$

would be required to produce this compression. When thus compressed, this portion of the gas would have a density greater than the equilibrium

density  $\rho$  by an amount  $\Delta\rho = \rho(\Delta V/V_S)$  and its pressure would be greater than  $P$  by an amount  $\Delta P = P(\Delta V/V_S)$ .

Therefore the potential energy corresponding to an increase of density of the part of the gas originally in volume  $V_S$ , from its equilibrium density  $\rho$  to a nonequilibrium density  $\rho + \Delta\rho$ , is  $(1/2)NkT \times (\Delta\rho/\rho)^2$ . For thermal fluctuations the mean potential energy is  $(1/2)kT$ , if the potential energy is a quadratic function of the variable  $\Delta\rho$  (as it is here). Consequently the mean-square fractional fluctuation of density of a portion of the gas containing  $N_S$  molecules (and occupying volume  $V_S$  at equilibrium) is

$$\langle(\Delta\rho/\rho)^2\rangle = 1/N_S \quad (15-6)$$

which is also equal to the mean-square fractional fluctuation of pressure,  $\langle(\Delta P/P)^2\rangle$ . Another derivation of this formula is given in Chapter 23.

We see that the smaller the fraction of the gas we look at (the smaller  $N_S$  is) the greater the fractional fluctuation of density and pressure is caused by thermal motion. If we watch a small group of molecules, their thermal motion will produce relatively large changes in their density. On the other hand if we include a large number of molecules in our sample, the large fluctuations in each small part of the sample will to a great extent cancel out, leaving a mean-square fractional fluctuation of the whole which is smaller the larger the number  $N_S$  of molecules in the sample. The root-mean-square fractional fluctuation of density or pressure of a portion of the gas is inversely proportional to the square root of the number of molecules sampled.

These fluctuations of density tend to scatter acoustical and electromagnetic waves as they travel through the gas. Indeed it is the scattering of light by the thermal fluctuations of the atmosphere which produces the blue of the sky. The fluctuations are independent of temperature, although at lower temperatures the  $N_S$  molecules occupy a smaller volume and thus the fluctuations are more "fine-grained" at lower temperatures.

Incidentally, we could attack the problem of density fluctuations by asking how many molecules happen to be in a given volume  $V_S$  at some instant, instead of asking what volume  $N_S$  molecules happen to occupy at a given instant, as we did here. The results will of course turn out the same, as will be shown in Chapter 23.

### Brownian Motion

The fluctuating motion of a small particle in a fluid, caused by the thermal fluctuations of pressure on the particle, is called Brownian motion. We have already seen [in Eq. (15-3)] that the mean square of

each of the velocity components of such motion is proportional to  $T$  and inversely proportional to the mass of the particle. The mean square of each of the position coordinates of the particle is not as simple to work out for the unbound particle as it was for the displacement of the mass on a spring, discussed in the preceding section.

In the case of the mass on the end of the spring, the displacement  $x$  from equilibrium is confined by the restoring force, the maximum displacement is determined by the energy possessed by the oscillator, and we can measure a mean-square displacement from equilibrium,  $\langle x^2 \rangle$ , by averaging the value of  $x^2$  over any relatively long interval of time (the longer the interval, the more accurate the result, of course). But the  $x$  component of displacement of a free particle in a fluid is not so limited; the only forces acting on the particle (if we can neglect the force of gravity) are the fluctuations of pressure of the fluid, causing the Brownian motion, and the viscous force of the fluid, which tends to retard the particle's motion. If we measure the  $x$  component of the particle's position (setting the initial position at the origin) we shall find that, although the direction of motion often reverses, the particle tends to drift away from the origin as time goes on, and eventually it will traverse the whole volume of fluid, just as any molecule of the fluid does. If we allow enough time, the particle is eventually likely to be anywhere in the volume. This is in correspondence with the Maxwell-Boltzmann distribution; the potential energy does not depend on  $x$  (neglecting gravity) so the probability density  $f$  is independent of  $x$ ; any value of  $x$  is equally likely in the end.

But this was not the problem at present. We assumed that the particle under observation started at  $x = 0$ ; it certainly isn't likely to drift far from the origin right away. Of course the average value of  $x$  is zero, since the particle is as likely to drift to the left as to the right. But the expected value of  $x^2$  must increase somehow with time. At  $t = 0$  the particle is certainly at the origin; as time goes on the particle may drift farther and farther away from  $x = 0$ , in either direction. We wish to compute the expected value of  $x^2$  as a function of time or, better still, to find the probability density of finding the particle at  $x$  after time  $t$ . Note that this probability is a conditional probability density; it is the probability of finding the  $x$  component of the position of the particle to be  $x$  at time  $t$  if the particle was at the origin at  $t = 0$ .

### Random Walk

We can obtain a better insight into this problem if we consider the random-walk problem of Eqs. (11-9) and (11-10). A crude model of one-dimensional Brownian motion can be constructed as follows. Suppose a particle moves along a line with a constant speed  $v$ . At the end of each successive time interval  $\tau$  it may change its direction of mo-

tion or not, the two alternatives being equally likely ( $p = 1/2$ ) and distributed at random. After  $N$  intervals (i.e., after time  $N\tau$ ) the chance that the particle is displaced an amount  $x_N = (2n - N)v\tau$  from its initial position is then the chance that during  $n$  of the  $N$  intervals it went to the right and during the other  $(N - n)$  intervals it went to the left; it covered a distance  $v\tau$  in each interval. According to Eq. (11-5) this probability is

$$P_n(N) = N!/(1/2)^N / n! (N - n)! \quad (15-7)$$

since  $p = 1/2$  for this case.

The expected value of  $x_N$  and its variance are then obtained from Eqs. (11-9) and (11-10) (for  $p = 1/2$ ):

$$\begin{aligned} \langle x \rangle &= \sum_{n=0}^N (2n - N)v\tau P_n(N) = 0 \\ \langle x^2 \rangle &= \sum_{n=0}^N (2n - N)^2 (v\tau)^2 P_n(N) = N(v\tau)^2 \end{aligned} \quad (15-8)$$

The expected value of the displacement is zero, since the particle is as likely to go in one direction as in the other. Its tendency to stray from the origin is measured by  $\langle x^2 \rangle$ , which increases linearly with the number of time intervals  $N$ , and thus increases linearly with time. If the particle, once started, continued always in the same direction ( $p = 0$  or  $1$ ) the value of  $\langle x^2 \rangle$  would be  $(Nv\tau)^2$ , increasing quadratically with time. But with the irregular, to-and-fro motion of the random walk,  $\langle x^2 \rangle$  increases only linearly with time. Figure 15-3 is a plot

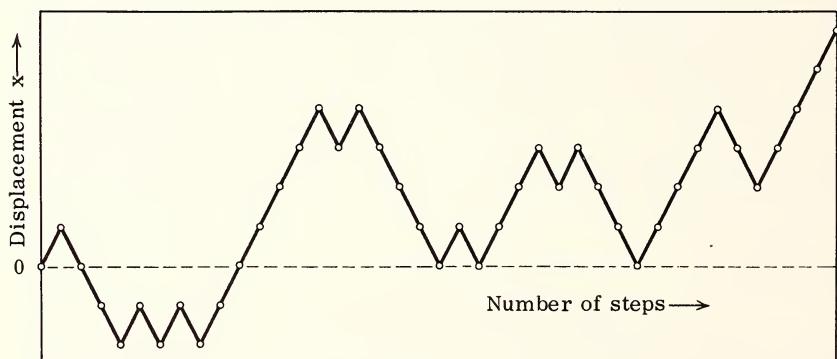


Fig. 15-3. Displacement for random walk. At each dot the "walker" flipped a coin to decide whether to step forward or backward.

of  $x$  as a function of  $t$  for a random walk as described here. We note the irregular character of the motion and the tendency to drift away from  $x = 0$ . Compare it with the curves of Fig. 15-1, and also with 15-2, for a mass with restoring force.

The limiting case of  $N$  very large and  $\tau$  very small is the case most nearly corresponding to Brownian motion. This limiting form was calculated at the end of Chapter 11. There we found that the probability distribution for displacement of the particle after  $N$  steps reduced, in the limit of  $N$  large, to the normal distribution of Eq. (11-17). The variance, in this case, as we just saw, is  $\sigma^2 = (v\tau)^2 N$  (for  $p = 1/2$ ) and the mean value of  $x$  is  $X = 0$ . Since the time required for the  $N$  steps is  $t = \tau N$  we can write the conditional probability density (that the particle will be at  $x$  at time  $t$  if it was at  $x = 0$  at  $t = 0$ ) as

$$F(x) = [1/(4\pi Dt)^{1/2}] e^{-x^2/4Dt}, \quad D = \frac{1}{2} v^2 \tau \quad (15-9)$$

so that the probability that the particle is between  $x$  and  $x + dx$  at time  $t$  is  $F(x) dx$ . We see that the "spread" of the distribution increases as  $t$  increases, as the particle drifts away from its initial position. The mean-square value of  $x$  is

$$\langle x^2 \rangle = \sigma^2 = (v\tau)^2 N = 2Dt$$

which increases linearly with time. Thus the question of the dependence of  $\langle x^2 \rangle$  on time, raised earlier in this section, is answered to the effect that  $\langle x^2 \rangle$  is proportional to  $t$ . The value of the proportionality constant  $2D$  for the actual Brownian motion of a particle in a fluid must now be determined.

### The Langevin Equation

To determine the value of constant  $D$  for a particle in a fluid we must study its equation of motion. As before, we study it in a single dimension first. The  $x$  component of the force on the particle can be separated into two parts, an average effect of the surrounding fluid plus a fluctuating part, caused by the pressure fluctuations of thermal motion of the fluid. The average effect of the fluid on the particle is a frictional force, caused by the fluid's viscosity. If the velocity of the particle in the  $x$  direction is  $\dot{x}$ , this average frictional force has an  $x$  component equal to  $M\beta \dot{x}$ , opposing the particle's motion, where the mechanical resistance to motion,  $M\beta$ , in a fluid of viscosity  $\eta$ , on a spherical particle of radius  $a$  is  $M\beta = 6\pi a \eta$  (Stoke's law). The fluctuating component of the force on the particle can simply be written as  $MA(t)$  (we write these functions with a factor  $M$ , the mass of the particle, so that  $M$  can be divided out in the resulting equation).

The equation of motion for the  $x$  component of the particle's position can thus be written as

$$M\ddot{x} = -M\beta\dot{x} + MA(t) \quad (15-10)$$

which is known as *Langevin's equation*. We note that  $\beta$  has the dimensions of reciprocal time. Multiplying the equation by  $x/M$  and using the identities

$$x\dot{x} = \frac{1}{2} \frac{d}{dt} (x^2) \quad \text{and} \quad x\ddot{x} = \frac{1}{2} \frac{d^2}{dt^2} (x^2) - (\dot{x})^2$$

we have

$$\frac{1}{2} \frac{d^2}{dt^2} (x^2) = -\beta \frac{d}{dt} (\dot{x}^2) + (x)^2 + xA(t)$$

This is an equation for one particular particle. If we had many identical particles in the fluid (or if we performed a sequence of similar observations on one particle) each particle would have different values of  $x$  and  $\dot{x}$  at the end of a given time  $t$ , because of the effects of the random force  $A(t)$ .

Suppose we average the effects of the fluctuations by averaging the terms of Eq. (15-10) over all similar particles. The term  $xA(t)$  will average out because both  $\langle x \rangle$  and  $\langle A \rangle$  are zero and the fluctuations of  $x$  and  $A$  are independent; the average value of  $\dot{x}^2$ , however, carries with it the mean effects of the fluctuating force  $A(t)$ . We showed in Eq. (15-3) that for a particle in thermal equilibrium at temperature  $T$ , its mean-square velocity component  $\langle \dot{x}^2 \rangle$  is equal to  $kT/M$ .

If we now symbolize the mean-square displacement  $\langle x^2 \rangle$  as  $s(t)$ , the average of the equation of motion written above turns out to be

$$\frac{1}{2} \ddot{s} = (kT/M) - \frac{1}{2} \beta \dot{s}$$

The solution of this equation is  $\dot{s} = (2kT/M\beta) - Ce^{-\beta t}$ . The transient exponential soon drops out, leaving for the steady-state solution  $\dot{s} = (2kT/M\beta)$  and thus

$$s = \langle x^2 \rangle = (2kT/M\beta)t \quad (15-11)$$

This result answers the question raised at the end of the previous section; the constant  $D = (1/2)(v^2\tau)$ , used there now turns out to be  $kT/M\beta$  and, for a spherical particle of radius  $a$  in a fluid of viscosity  $\eta$ , constant  $D$  is equal to  $kT/6\pi a\eta$  from Stoke's law.

The innocuous looking result shown in Eq. (15-11) enabled Perrin and others first to measure Avogadro's number  $N_0$  and thus, in a sense, first to make contact with atomic dimensions. They were able to measure  $N_0$  in terms of purely macroscopic constants, plus obser-

vations of Brownian motion. A spherical particle was used, of known radius  $a$ , so that Stoke's law applied. The viscosity of the fluid in which the particle was immersed was measured as well as the temperature  $T$  of the fluid. The value of the gas constant  $R$  was known but at the time neither the value of the Boltzmann constant  $k$  nor the value of  $N_0 = R/k$  was known. The  $x$  coordinates of the particle in the fluid were measured at the ends of successive intervals of time of length  $t$ ;  $x_0$  at  $t = 0$ ,  $x_1$  at  $t$ ,  $x_2$  at  $2t$ , and so on, and the average of the set of values  $(x_{n+1} - x_n)^2$  was computed, which is equivalent to the  $\langle x^2 \rangle$  of Eq. (15-11).

By making the measurements for several different values of the time interval  $t$ , it was verified that  $\langle x^2 \rangle$  does indeed equal  $2Dt$ , and the value of  $D$  was determined. The value of Avogadro's number,

$$N_0 = R/k = RT/6\pi a \eta D$$

can thus be computed. By this method a value of  $N_0$  was obtained which checks within about 5 parts in a thousand with values later obtained by more direct methods. Of course very small spheres had to be used, to make  $\langle x^2 \rangle$  as large as possible, and careful observations with a microscope were made to determine the successive  $x_n$ 's. Perrin used spheres of radius  $2 \times 10^{-5}$  cm and time intervals from a few seconds to a minute or more.

### The Fokker-Planck Equations

Brownian motion is simply the fine details of the process of diffusion. If there were initially a concentration of particles in one region of the fluid, as time went on these particles would diffuse, by Brownian motion, to all parts of the fluid. The diffusion constant for a particle in a fluid is  $D = kT/M\beta$ , which is to be compared with Eq. (14-8) for the  $D$  for a molecule; in the molecular case  $\beta$  is evidently equal to  $1/t_c$ , whereas for a larger, spherical particle  $\beta = 6\pi a \eta / M$ . The mean concentration of the diffusing particles must satisfy a diffusion equation of the type given in Eq. (14-10).

This means that there is a close connection between the results of this chapter and those of the section on diffusion. Whether the diffusing entity undergoing Brownian motion is a molecule of the gas or a dust particle in the gas, the probability density of its presence at the point in space given by the vector  $\mathbf{r}$  at time  $t$ , if it starts at  $\mathbf{r}_0$  at  $t = 0$ , is given by the three-dimensional generalization of Eq. (15-9),

$$f_r(\mathbf{r}, t) = (4\pi Dt)^{-3/2} \exp\left(-\frac{|\mathbf{r} - \mathbf{r}_0|^2}{4Dt}\right); \quad D = kT/M\beta \quad (15-12)$$

which is a solution of the diffusion equation (14-10). The value of  $\beta$

appropriate for the particle under study must be used, of course.

The distribution function for diffusion by Brownian motion of Eq. (15-12) and the diffusion equation (14-10) that it satisfies can thus be derived by the methods of the previous chapter or else by those of this chapter. For example, it is possible to generalize the Langevin equation (15-10) and manipulate it to obtain the diffusion equation. Also, by either method, it can be shown that when an external force  $\mathbf{F}$  acts on the diffusing particle (such as the force of gravity), the diffusion equation has the more general form

$$\partial f_r / \partial t = \text{div} [D \text{ grad } f_r - (\mathbf{F}/M\beta) f_r] \quad (15-13)$$

When  $f_r$  is the density of diffusing substance (molecules or heat, for example), Eq. (15-13), or its simple version (14-10) for  $\mathbf{F} = 0$ , is called the diffusion equation. When  $f_r$  is the distribution function for a particle undergoing Brownian motion and the equation is considered to be a first approximation to a generalized Langevin equation, then Eq. (15-13) is called a *Fokker-Planck equation*. The solutions behave the same in either case, of course. The solution of (15-13) for  $\mathbf{F} = 0$  and for the particle starting at  $\mathbf{r} = \mathbf{r}_0$  when  $t = 0$  is Eq. (15-12). From it can be derived all the characteristics of Brownian motion in regard to the possible position of the particle at time  $t$ .

A Fokker-Planck equation can also be obtained for the distribution in momentum,  $f_p(p, t)$  of the particle. It is

$$\partial f_p / \partial t = \beta \text{ div}_p (MkT \text{ grad}_p f_p + f_p p) \quad (15-14)$$

For a particle that is started at  $t = 0$  with a momentum  $\mathbf{p} = \mathbf{p}_0$ , the solution of this equation, which is the probability density of the particle in momentum space, is

$$f_p(p, t) = [2\pi MkT(1 - e^{-2\beta t})]^{-3/2} \exp \left[ \frac{-|p - p_0 e^{-\beta t}|^2}{2MkT(1 - e^{-2\beta t})} \right] \quad (15-15)$$

This interesting solution shows that the expected momentum of the particle at time  $t$  is  $p_0 e^{-\beta t}$  [compare with the discussion of Eq. (11-17)], which is the momentum of a particle started with a momentum  $\mathbf{p}_0$  and subjected to a frictional retarding force  $-\beta p$ . As time goes on, the effect of the fluctuations "spreads out" the distribution in momentum; the variance of the momentum (i.e., its mean-square deviation from  $p_0 e^{-\beta t}$ ) is  $kT(1 - e^{-2\beta t})$ , starting as zero when  $t = 0$ , when we are certain that the particle's momentum is  $\mathbf{p}_0$ , and approaching asymptotically the value  $kT$ , which is typical of the Maxwell distribution. Thus Eq. (15-15) shows how an originally nonequilibrium distribution for a particle (or a molecule) in a fluid can change with time into the Maxwell distribution typical of an equilibrium state. Constant

$\beta$ , which equals  $6\pi a\eta/M$  for a spherical particle or  $1/t_c$  for a molecule in a gas, is thus equal to the reciprocal of the relaxation time for the distribution, which relates directly to the discussion of Eq. (14-2).

Of course the most general distribution function would be  $f(r, p, t)$ , giving the particle's distribution in both position and momentum at time  $t$  after initial observation. The equation for this  $f$  is, not surprisingly, closely related to the Boltzmann equation (13-2). It can be shown to be

$$\frac{\partial f}{\partial t} + \frac{p}{M} \cdot \text{grad}_r f + \mathbf{F} \cdot \text{grad}_p f = \beta \text{div}_p [MkT \text{grad}_p f + pf] \quad (15-16)$$

The derivation of this equation, particularly of the right-hand side of it, involves a generalization of the arguments used in deriving Eq. (15-11). This right-hand side is another approximation to the collision term  $Q$  of Eq. (13-2).



# III

## STATISTICAL MECHANICS



# 16

## Ensembles and Distribution Functions

It is now time to introduce the final generalization, to present a theoretical model that includes all the special cases we have been considering heretofore. If we had been expounding our subject with mathematical logic we would have started at this point, presenting first the most general assumptions and definitions as postulates, working out the special cases as theorems following from the postulates, and only at the end demonstrating that the predictions, implicit in the theorems, actually do correspond to the "real world," as measured by experiment. We have not followed this procedure, for several reasons.

In the first place, most people find it easier to understand a new subject, particularly one as complex as statistical physics, by progressing from the particular to the general, from the familiar to the abstract.

A more important reason, however, is that physics itself has developed in a nonlogical way. Experiments first provide us with data on many particular cases, at first logically unconnected with each other, which have to be learned as a set of disparate facts. Then it is found that a group of these facts can be considered to be special cases of a "theory," an assumed relationship between defined quantities (energy, entropy, and the like) which will reproduce the experimental facts when the theory is appropriately specialized. This theory suggests more experiments, which may force modifications of the theory and may suggest further generalizations until, finally, someone shows that the whole subject can be "understood" as the logical consequences of a few basic postulates.

At this point the subject comes to resemble a branch of mathematics, with its postulates and its theorems logically deduced therefrom. But the similarity is superficial, for in physics the experimental facts are basic and the theoretical structure is erected to make it

easier to "understand" the facts and to suggest ways of obtaining new facts. A logically connected theory turns out to be more convenient to remember than are vast arrays of unconnected data. This convenience, however, should not persuade us to accord the theory more validity than should inhere in a mnemonic device. We must not expect, for example, that the postulates and definitions should somehow be "the most reasonable" or "the logically inevitable" ones. They have been chosen for the simple, utilitarian reason that a logical structure reared on them can be made to correspond to the experimental facts. Thus the presentation of a branch of physics in "logical" form tends to exaggerate the importance and inevitability of the theoretical assumptions, and to make us forget that the experimental data are the only truly stable parts of the whole.

This danger, of ascribing a false sense of inevitability to the theory, is somewhat greater with statistical physics than with other branches of classical physics, because the connection between experiment and basic theory is here more indirect than usual. In classical mechanics the experimental verification of Newton's laws can be fairly direct; and the relationship between Faraday's and Ampere's experiments and Maxwell's equations of electromagnetic theory is clearcut. In thermodynamics, the experiments of Rumford, relating work and heat, bear a direct relationship to the first law, but the experimental verification of the second law is indirect and negative. Furthermore, the more accurate "proofs" that the Maxwell-Boltzmann distribution is valid for molecules in a gas, are experimentally circuitous. And finally, as we shall see later, there is no experiment, analogous to those of Faraday or Ampere, which directly verifies any of the basic assumptions of statistical mechanics; their validity must be proved piecemeal and inferentially. In the end, of course, the proofs are convincing from their very number and breadth of application.

However we have now reached a point in our exposition where the basic theory must be presented, and it is necessary to follow the pattern of mathematical logic for a time. Our definitions and postulates are bound to sound arbitrary until we have completed the demonstration that the theory does indeed correspond to a wide variety of observed facts. But we must keep in mind that they have been chosen solely to obtain this correspondence with observation, not because they "sound reasonable" or satisfy some philosophical "first principles."

### Distribution Functions in Phase Space

In Chapters 13 and 15 we discussed distribution functions for molecules, and also for multimolecular particles in a gas. In statistical

mechanics we carry this generalization to its logical conclusion, and deal with distribution functions for complete thermodynamic systems. A particular microstate of such a system (a gas of  $N$  particles, for example) can be specified by choosing values for the  $6N$  position and momentum coordinates; the distribution function is the probability density that the system has these coordinate values. Geometrically speaking, an elementary region in this  $6N$ -dimensional phase space represents a microstate of the system; the point representing the system passes through all the microstates allowed by its thermodynamic state; the fraction of time spent by the system point in a particular region of phase space is proportional to the distribution function corresponding to the thermodynamic state. In other words, a choice of a particular thermodynamic state (a macrostate) is equivalent to a choice of a particular distribution function, and vice versa. The task of statistical mechanics is to devise methods for finding distribution functions which correspond to specific macrostates.

According to classical mechanics, the distribution function for a system with  $\phi = 3N$  degrees of freedom is

$$f(q, p) = f(q_1, q_2, \dots, q_\phi, p_1, p_2, \dots, p_\phi)$$

where the  $q$ 's are the coordinates and the  $p$ 's the conjugate momenta [see Eq. (13-9)] which specify the configuration of the system as a whole. Then  $f(q, p) dV_q dV_p$  [where  $dV_q = h_1 dq_1 h_2 dq_2 \dots h_\phi dq_\phi$  and  $dV_p = (dp_1/h_1)(dp_2/h_2) \dots (dp_\phi/h_\phi)$ ] is the probability that the system point is within the element of phase space  $dV_q dV_p$  at position  $q_1, \dots, p_\phi$ , at any arbitrarily chosen instant of time.

More generally, the distribution function represents the probability density, not for one system, but for a collection of similar systems. Imagine a large number of identical systems, all in the same thermodynamic state but, of course, each of them in different possible microstates. This collection of sample systems is called an *ensemble* of systems, the ensemble corresponding to the specified macrostate. Different ensembles, representing different thermodynamic states, have different populations of microstates. The distribution function for the macrostate measures the relative number of systems in the ensemble which are in a given microstate at any instant. Thus it is a generalization of the distribution function of Eq. (13-10), which was for an ensemble of molecules.

Each system in the ensemble has  $\phi$  coordinates  $q_i$  and  $\phi$  momenta  $p_i$ , and a Hamiltonian function  $H(q, p)$ , which is the total energy of the system, expressed in terms of the  $q$ 's and  $p$ 's. The values of these  $q$ 's and  $p$ 's at a given instant determine the position of its system point in phase space. The motion of the system point in phase space is determined by the equations of motion of the system. These

can be expressed in many forms, each of which are discussed in books on classical dynamics. The form which is most useful for us at present is the set of Hamilton's equations,

$$\dot{q}_i = \partial H / \partial p_i; \quad \dot{p}_i = -\partial H / \partial q_i, \quad i = 1, 2, \dots, \phi \quad (16-1)$$

the first set relating the velocities  $\dot{q}_i$  to the momenta and the second set relating the force components to the rates of change of the momenta (the "velocity components" in momentum space).

The ensemble of systems can thus be represented as a swarm of system points in phase space, each point moving in accordance with Eqs. (16-1); the velocity of the point in phase space is proportional to the gradient of  $H$ . The density of points in any region of phase space is proportional to the value of the distribution function  $f(q, p)$  in that region.

### Liouville's Theorem

If the thermodynamic state is not an equilibrium state,  $f$  will be a function of time. If the state is an equilibrium state, the density of system points in any specified region of phase space will be constant; as many system points will enter the region per unit of time as will leave it. The swarm of system points has some similarity to the swarm of particles in a gas. The differences are important, however. The system points are moving in a  $2\phi$ -dimensional phase space, not real space; also the system points do not collide. In fact the system points do not interact at all, for each system point represents a different system of the ensemble, and the separate systems cannot interact since they are but samples in an imaginary array of systems, assembled to represent a particular macrostate. Each individual system point, for example, may represent a whole gas of  $N$  atoms, or a crystal lattice, depending on the situation the ensemble has been chosen to represent. There can be no physical interaction between the individual sample systems.

This means that the Boltzmann equation for the change of  $f$  with time, the generalization of Eq. (13-2) to the ensemble, has no collision term  $Q$ . The equation,

$$\frac{\partial f}{\partial t} + \sum_{i=1}^{\phi} \frac{\partial}{\partial q_i} (\dot{q}_i f) + \sum_{i=1}^{\phi} \frac{\partial}{\partial p_i} (\dot{p}_i f) = 0 \quad (16-2)$$

is simply the equation of continuity in phase space, and represents the fact that, as the swarm of system points moves about in phase space, no system point either appears or disappears.

Since each system in the ensemble obeys Hamilton's equations (16-1), this equation of continuity becomes

$$\frac{\partial f}{\partial t} + \sum_{i=1}^{\phi} \left[ \frac{\partial}{\partial q_i} \left( f \frac{\partial H}{\partial p_i} \right) - \frac{\partial}{\partial p_i} \left( f \frac{\partial H}{\partial q_i} \right) \right] = 0$$

and since

$$\frac{\partial H}{\partial p_i} \frac{\partial f}{\partial q_i} = \frac{\partial}{\partial q_i} \left( f \frac{\partial H}{\partial p_i} \right) - f \frac{\partial^2 H}{\partial q_i \partial p_i}$$

etc., we have

$$\frac{\partial f}{\partial t} + \sum_{i=1}^{\phi} \left[ \left( \frac{\partial H}{\partial p_i} \frac{\partial f}{\partial q_i} \right) - \left( \frac{\partial H}{\partial q_i} \frac{\partial f}{\partial p_i} \right) \right] = 0$$

$$\frac{\partial f}{\partial t} \equiv \frac{\partial f}{\partial t} + \mathbf{v} \cdot \mathbf{grad} f \equiv \frac{\partial f}{\partial t} + \sum_{i=1}^{\phi} \dot{q}_i \frac{\partial f}{\partial q_i} + \sum_{i=1}^{\phi} \dot{p}_i \frac{\partial f}{\partial p_i} = 0 \quad (16-3)$$

where  $df/dt$  is the change in the distribution function  $f$  in a coordinate system which moves with the system points. Because of the relationship between  $\dot{q}$  and  $p$ ,  $\dot{p}$  and  $q$ , inherent in Hamilton's equations, the density of system points near a given system point of the ensemble remains constant as the swarm moves about. If, at  $t = 0$ , the swarm has a high density in a localized region of phase space, this concentration of system points moves about as time goes on but it does not disperse; it keeps its original high density. This result is known as *Liouville's theorem*.

We can use Liouville's theorem to devise distribution functions which are independent of time, i.e., which represent equilibrium macrostates. For example, if  $f$  had the same value everywhere in phase space, it would be independent of time; as a part of the swarm moved away from a given region of phase space a different part of the swarm would move in and, since all parts of the swarm have (and keep) the same density, the density in a given region would not change. We can also devise less trivial stationary distributions, for the path traversed by any one system point does not cover all of phase space; it confines itself to the hypersurface on which the Hamiltonian function  $H(q, p)$  is constant; an isolated system cannot change its total

energy. Therefore if the distribution function is the same for all regions of phase space for which  $H$  is the same (i.e., if  $f$  is a function of  $H$  alone) the density of system points in a given region cannot change as the points move along their constant- $H$  paths.

We shall deal with several different types of distribution functions, corresponding to different specifications regarding the thermodynamic state of the system. The simplest one is for  $f$  to be zero everywhere in phase space except on the hypersurface corresponding to  $H(q,p) = E$ , a constant; the ensemble corresponding to this is called a *microcanonical ensemble*. A more useful case is for  $f$  to be proportional to  $\exp[-H(q,p)/kT]$ , corresponding to what is called the *canonical ensemble*. Other ensembles, with  $f$ 's which are more-complicated functions of  $H$ , will also prove to be useful. But, in order for any of them to represent actual thermodynamic macrostates, we must assume a relationship between the distribution function  $f$  for an ensemble and the corresponding thermodynamic properties of the macrostate which the ensemble is supposed to represent. The appropriate relationship turns out to be between the entropy of the macrostate and the distribution function of the corresponding ensemble.

### Quantum States and Phase Space

Before we state the basic postulate of statistical mechanics, relating entropy and ensembles, we should discuss the modifications which quantum theory makes in our definitions. In some respects the change is in the direction of simplification, the summation over denumerable quantum numbers being substituted for integration over continuous coordinates in phase space. Instead of Hamilton's equations (16-1), there is a Shrödinger equation for a wave function  $\Psi(q_1, q_2, \dots, q_\phi)$  and an allowed value  $E$  of energy of the system, both of which depend on the  $\phi$  quantum numbers  $\nu_1, \nu_2, \dots, \nu_\phi$ , which designate the quantum state of the system.

For example, if the system consists of  $N$  particles in the simple crystal model of Eq. (13-2), the classical Hamiltonian for the whole system can be written as

$$H(q,p) = \sum_{j=1}^{\phi} [1/2m)p_j^2 + (m\omega^2/2)q_j^2] \quad (\phi = 3N) \quad (16-4)$$

where  $q_{3i-2} = x_i$ ,  $q_{3i-1} = y_i$ ,  $q_{3i} = z_i$ ,  $p_{3i-2} = p_{xi}$ , etc. Hamilton's equations become

$$\dot{q}_i = (1/m)p_i; \quad \dot{p}_i = m\omega^2q_i \quad (16-5)$$

and Schrödinger's equation for the system is  $H\Psi = E\Psi$ , where each

$p_i$  in the  $H$  is changed to  $(\hbar/i)(\partial/\partial q_i)$ . For (16-4) it is

$$\sum_{i=1}^{\phi} \left[ \frac{1}{2} m \omega^2 q_i^2 - \frac{\hbar^2}{2m} \frac{\partial^2}{\partial q_i^2} \right] \Psi = E \Psi \quad (16-6)$$

where Planck's constant  $h$  is equal to  $2\pi\hbar$ . Solution of this equation, subject to the requirement that  $\Psi$  is finite everywhere, results in the following allowed values of the energy  $E$ ,

$$E_{\nu_1, \nu_2, \dots, \nu_\phi} = \sum_{i=1}^{\phi} \hbar \omega \left( \nu_i + \frac{1}{2} \right) \quad (16-7)$$

The distribution function for an equilibrium state is a function of  $E_{\nu_1, \dots, \nu_\phi}$  and thus is a function of the  $\phi$  quantum numbers  $\nu_1, \nu_2, \dots, \nu_\phi$ , instead of being a function of  $H(q, p)$  and thus a function of the  $2\phi$  continuous variables  $q_1, q_2, \dots, q_\phi, p_1, \dots, p_\phi$ , as it was in classical mechanics. Function  $f(\nu_1, \dots, \nu_\phi)$  is the probability that the system is in the quantum state characterized by the quantum numbers  $\nu_1, \dots, \nu_\phi$ ; as contrasted with the probability  $f(q_1, \dots, p_\phi) dq_1 dp_1$  for phase space. These statements apply to any system, not just to the simple crystal model. The quantum state for any system with  $\phi$  degrees of freedom, no matter what conservative forces its particles are subjected to, is characterized by  $\phi$  quantum numbers  $\nu_1, \dots, \nu_\phi$ . To simplify notation we shall often write the single symbol  $\nu$  instead of the  $\phi$  individual numbers  $\nu_1, \dots, \nu_\phi$ , just as we write  $q$  for  $q_1, \dots, q_\phi$ , etc. Thus  $f_\nu$  is the probability that the system is in the quantum state  $\nu \equiv \nu_1, \dots, \nu_\phi$ . The sum  $\sum_\nu f_\nu$  over all allowed states of the system must be unity, of course.

We thus have two alternative ways of expressing the microstates of the system, and thus of writing the distribution function. The quantum way, saying that each quantum state of the system is a separate and distinct microstate, is the correct way, but it sometimes leads to computational difficulties. The classical way, of representing a microstate as a region of phase space, is only an approximate representation, good for large energies; but when valid it is often easier to handle mathematically. The quantitative relationship between these two ways is obtained by working out the volume of phase space "occupied" by one quantum state of the system.

The connection between the classical coordinates  $q_i$  and momenta  $p_i$  and the quantum state is provided by the Heisenberg uncertainty

principle,  $\Delta q_i \cdot \Delta p_i \geq h$ . A restatement of this is that, in the phase space of one degree of freedom, one quantum state occupies a "volume"  $\Delta q_i \Delta p_i$  equal to  $h$ . For example, the one-dimensional harmonic oscillator has a Hamiltonian  $H_i = (1/2m)p_i^2 + (m\omega^2/2)q_i^2$ . When in the quantum state  $\nu_i$ , with energy  $\hbar\omega(\nu_i + 1/2) = (h\omega/2\pi)(\nu_i + 1/2)$ , its phase-space orbit is an ellipse in phase space, with semiminor axis  $q_m = [(h/\pi m\omega)(\nu_i + 1/2)]^{1/2}$  along  $q_i$  and semimajor axis  $p_m = [(hm\omega/\pi)(\nu_i + 1/2)]^{1/2}$ , which encloses an area

$$A(\nu_i) = \pi p_m q_m = h\left(\nu_i + \frac{1}{2}\right)$$

The area between successive ellipses, for successive values of  $\nu_i$ , is the area "occupied" by one quantum state. We see that  $A(\nu_i + 1) - A(\nu_i) = h$ , as stated above.

Thus a volume element  $dq_i dp_i$  corresponds, on the average, to  $(dq_i dp_i/h)$  quantum states. Similarly, for the whole system, with  $\phi$  degrees of freedom, the volume element  $dV_q dV_p = dq_1 \dots dp_\phi$  will correspond, on the average, to  $(dV_q dV_p/h^\phi)$  quantum states. Thus the correspondence between volume of phase space and number of quantum states is

$$\text{No. of microstates} = (1/h^\phi)(\text{vol. of phase space}) \quad (16-8)$$

when the system has  $\phi$  degrees of freedom.

When the volume of phase space occupied by the swarm of system points, representing a particular ensemble, is very large compared to  $h^\phi$ , the classical representation, in terms of the continuous variables  $q_1, \dots, p_\phi$  can be safely used. But when the volume occupied by the swarm is not large compared to  $h^\phi$ , the classical representation is not likely to be valid and the quantum representation is needed [see Eq. (19-8) et seq.].

# 17

## Entropy and Ensembles

As pointed out in the preceding chapter, we are presenting statistical mechanics in "logical" order, with definitions and basic postulates first, theorems and connections with experiment later. The last chapter was devoted to definitions. Each thermodynamic macrostate of a system may be visualized as an ensemble of systems in a variety of microstates, or may be represented quantitatively in terms of a distribution function, which is the probability  $f_\nu$  that a system chosen from the ensemble is in the quantum state  $\nu = \nu_1, \nu_2, \dots, \nu_\phi$  or is the probability density  $f(q, p)$  that the system point has the coordinates  $q_1, q_2, \dots, p_\phi$  in phase space, if the macrostate is such that classical mechanics is valid. In this chapter we shall introduce the essential postulates.

### Entropy and Information

The basic postulate, relating the distribution function  $f_\nu$  to the thermodynamic properties of the macrostate which the ensemble represents, was first stated by Boltzmann and restated in more general form by Planck. In the form appropriate for our present discussion it relates the entropy  $S$  of the system to the distribution function  $f_\nu$  by the equation

$$S = -k \sum_\nu f_\nu \ln(f_\nu); \quad \sum_\nu f_\nu = 1 \quad (17-1)$$

where  $k$  is the Boltzmann constant and where the summation is over all the quantum states present in the ensemble (i.e., for which  $f_\nu$  differs from zero).

This formula satisfies our earlier statements that  $S$  is a measure of the degree of disorder of the system [see discussion preceding Eq. (6-14)]. A system that is certainly in its single, lowest quantum state is one in perfect order, so its entropy should be zero. Such a

system would have the  $f_\nu$  for the lowest quantum state equal to unity and all the other  $f$ 's would be zero. Since  $\ln(1) = 0$  and  $x \ln(x) \rightarrow 0$  as  $x \rightarrow 0$ , the sum on the right-hand side of Eq. (17-1) is zero for this case. On the other hand, a disorderly system would be likely to be in any of a number of different quantum states; the larger the number of states it might occupy the greater the disorder. If  $f_\nu = 1/N$  for  $N$  different microstates (label them  $\nu = 1, 2, \dots, N$ ) and  $f_\nu$  is zero for all other states then

$$S = -k \sum_{\nu=1}^N (1/N) \ln(1/N) = k \ln N$$

which increases as  $N$  increases. Thus Eq. (17-1) satisfies our pre-conceptions of the way entropy should behave. It also provides an opportunity to be more exact in regard to the measurement of disorder.

Disorder, in the sense we have been using it, implies a lack of information regarding the exact state of the system. A disordered system is one about which we lack complete information. Equation (17-1) is the starting point for Shannon's development of information theory. It will be useful to sketch a part of this development, for it will cast further light on the meaning of entropy, as postulated in Eq. (17-1).

### Information Theory

A gasoline gauge, with a pointer and scale, gives us more information about the state of the gasoline tank of an automobile than does a red light, which lights if the tank is nearly empty. How much more? Information comes to us in messages and to convey information each message must tell us something new, i.e., something not completely expected. Quantitatively, if there are  $N$  possible messages that could be received, and if the chance that the  $i$ -th one will be sent is  $f_i$ , then the information  $I$  that would be gained if message  $i$  were received must be a function  $I(f_i)$ , which increases as  $1/f_i$  increases. The less likely the message, the greater the information conveyed if the message is sent.

We can soon determine what function  $I(f_i)$  must be, for we require that information be additive; if two messages are received and if the messages are independent, then the information gained should be the sum of the  $I$ 's for each individual message. If the probability of message  $i$  be  $f_i$  and that for  $j$  be  $f_j$  then, if the two are independent, Eq. (11-3) requires that the probability that both messages happen to be sent is  $f_i f_j$ . The additive requirement for information then requires that

$$I(f_i f_j) = I(f_i) + I(f_j)$$

and this, in turn, requires that function  $I$  be a logarithmic function of  $f_i$ ,

$$I(f_i) = -C \ln f_i$$

where  $C$  is a constant. This is the basic definition of information theory. Since  $0 \leq f_i \leq 1$ ,  $I$  is positive and increases as  $1/f_i$  increases, as required.

The definition satisfies our preconceptions of how information behaves. For example, if we receive a message that is completely expected (i.e., its a priori probability is unity) we receive no information and  $I$  is zero. The less likely is the message (the smaller is  $f_i$ ) the greater the amount of information conveyed if it does get sent. The chance that the warning light of the gasoline gauge is off (showing that the tank is not nearly empty) is 0.9, say, so the information conveyed by the fact that the light is not lit is a small amount, equal to  $-C \ln 0.9 \approx 0.1C$ . On the other hand if the gauge has a pointer and five marks, each of which represents an equally likely state of the tank, then the information conveyed by a glance at the gauge is  $C \ln 5 \approx 1.6C$ , roughly 16 times the information conveyed by the unlit warning light (the information conveyed by a lit warning light, however, is  $C \ln 10 \approx 2.3C$ , a still larger amount).

To see how these definitions relate to our discussion of information and disordered systems, let us return to an ensemble, corresponding to some thermodynamic state, with its distribution function  $f_\nu$ . If we wish to find out exactly what microstate the system happens to be in at any instant, we would subject it to detailed measurement designed to tell us. The results of the measurement would be a kind of message to us, giving us information. If the measurements happened to show that the system is in microstate  $\nu$ , the information gained by the measurement would be  $-C \ln f_\nu$ , for  $f_\nu$  is the probability that the system would happen to be in state  $\nu$ . We of course cannot make a detailed enough examination to determine exactly what microstate a complicated system should happen to be in, nor would we wish to do so even if we could. But we can use the expected amount of information we *would* obtain, if we made the examination, as a measure of our present *lack* of knowledge of the system, i.e., of the system's disorder.

The expected amount of information we would obtain if we did examine the system in detail is the weighted mean of  $-C \ln f_\nu$  over all quantum states  $\nu$  in the ensemble, the weighting factor being the probability  $f_\nu$  of receiving the message that the system is in state  $\nu$ . This is the sum  $-C \sum_\nu f_\nu \ln f_\nu = (C/k)S$ , according to Eq. (17-1).

Thus the entropy  $S$  is proportional to our *lack of detailed information* regarding the system when it is in the thermodynamic state

corresponding to the distribution function  $f_\nu$ . Here again Eq. (17-1) corresponds to our preconceptions regarding the entropy  $S$  of a system.

### Entropy for Equilibrium States

But we do not wish to use postulate (17-1) to compute the entropy of a thermodynamic state when we know its distribution function; we wish to use Eq. (17-1) to find the distribution function for specified thermodynamic states, particularly those for equilibrium states. In order to do this we utilize a form of the second law. We noted in our initial discussion of entropy [see Eq. (6-5)] that in an isolated system  $S$  tends to increase until, at equilibrium, it is as large as it can be, subject to the restrictions on the system. If the sum of Eq. (17-1) is to correspond to the entropy, defined by the second law, it too must be a maximum, subject to restrictions, for an equilibrium state. These requirements should serve to determine the form of the distribution function, just as the wave equation, plus boundary conditions, determines the form of a vibrating string.

To show how this works, suppose we at first impose no restrictions on  $f_\nu$ , except that  $\sum_\nu f_\nu = 1$  and that the number of microstates in the ensemble represented by  $f_\nu$  is finite (so that the quantum number  $\nu$  can take on the values  $1, 2, \dots, W$ , where  $W$  is a finite integer). Then our problem is to determine the value of each  $f_\nu$  so that

$$S = -k \sum_{\nu=1}^W f_\nu \ln f_\nu \text{ is maximum, subject to } \sum_{\nu=1}^W f_\nu = 1 \quad (17-2)$$

This is a fairly simple problem in the calculus of variations, which can be solved by the use of Lagrange multipliers. But to show how Lagrange multipliers work, we shall first solve the problem by straightforward calculus.

The requirement that  $\sum f_\nu = 1$  means that only  $W-1$  of the  $f$ 's can be varied independently. One of the  $f$ 's, for example  $f_W$ , depends on the others through the equation  $f_W = 1 - \sum_{\nu=1}^{W-1} f_\nu$ . Now  $S$  is a

symmetric function of all the  $f$ 's, so we can write it  $S(f_1, f_2, \dots, f_W)$ , where we can substitute for  $f_W$  in terms of the others. In order that  $S$  be maximum we should have the partial derivative of  $S$  with respect to each independent  $f$  be zero. Taking into account the fact that  $f_W$  depends on all the other  $f$ 's, these equations become

$$(\partial S / \partial f_1) + (\partial f_W / \partial f_1)(\partial S / \partial f_W) = (\partial S / \partial f_1) - (\partial S / \partial f_W) = 0$$

$$(\partial S / \partial f_2) - (\partial S / \partial f_W) = 0$$

(17-3)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

$$(\partial S / \partial f_W - 1) - (\partial S / \partial f_W) = 0$$

The values of the  $f$ 's which satisfy these equations, plus the equation  $\sum f_\nu = 1$ , are those for which Eq. (17-2) is satisfied. For these values of the  $f$ 's, the partial derivative  $\partial S/\partial f_w$  will have some value; call it  $-\alpha_0$ . Then we can write Eqs. (17-3) in the more symmetric form

$$\left( \frac{\partial S}{\partial f_1} \right) + \alpha_0 = 0 \quad \left( \frac{\partial S}{\partial f_2} \right) + \alpha_0 = 0 \quad \dots \quad \left( \frac{\partial S}{\partial f_W} \right) + \alpha_0 = 0$$

However this is just the set of equations we would have obtained if, instead of the requirement that  $S(f_1, \dots, f_W)$  be maximum, subject to  $\sum f_\nu = 1$  of Eq. (17-2), we had instead used the requirement that

$$S(f_1, \dots, f_W) + \alpha_0 \sum_{\nu=1}^W f_{\nu} \text{ be maximum,} \quad (17-4)$$

$\alpha_0$  determined so that  $\sum_{\nu=1}^W f_{\nu} = 1$

Constant  $\alpha_0$  is a Lagrange multiplier.

Let us now solve the set of equations (17-4), inserting Eq. (17-1) for  $S$ . We set each of the partials of  $S + \alpha_0 \sum f_\nu$  equal to zero. For example, the partial with respect to  $f_k$  is

$$0 = \frac{\partial}{\partial f_K} \left[ \alpha_0 \sum_{\nu=1}^W f_{\nu} - k \sum_{\nu=1}^W f_{\nu} \ln f_{\nu} \right] = \alpha_0 - k \ln f_K - k$$

or

$$f_k = \exp [(\alpha_0/k) - 1]$$

The solution indicates that all the  $f$ 's are equal, since neither  $\alpha_0$  nor  $k$  depend on  $\kappa$ . The determination of the value of  $\alpha_0$ , and thus of the magnitude of  $f_\nu$ , comes from the requirement  $\sum f_\nu = 1$ :  $f_\nu = (1/W)$ , so that

$$S = -k \sum_{\nu=1}^W \left(\frac{1}{W}\right) \ln\left(\frac{1}{W}\right) = k \ln W \quad (17-5)$$

For a system restricted to a finite number  $W$  of microstates, and with no other restrictions, the state of maximum entropy is that for which the system is equally likely to be in any of the  $W$  microstates, and the corresponding maximal value of the entropy is  $k$  times the natural logarithm of the number  $W$  (which is sometimes called the *statistical weight* of the equilibrium macrostate).

### Application to a Perfect Gas

To show how much is inherent in these abstract-sounding results, we apply them to a gas of  $N$  point particles, confined in a container of volume  $V$ . To say that the system is confined to a finite number of quantum states is equivalent, classically, to saying that the system point is confined to a finite volume in phase space. In fact, from Eq. (16-6), the volume of phase space  $\Omega$ , within which the system point is to be found, is  $\Omega = h^{3N}W$ , where  $\phi = 3N$  is the number of degrees of freedom of the gas of  $N$  particles. And, from Eq. (17-5) we see that the system point is *equally likely* to be anywhere within this volume  $\Omega$ . Thus, classically the analogue of Eq. (17-5) is

$$f(q, p) = 1/\Omega; \quad S = k \ln(\Omega/h^{3N}) \quad (17-6)$$

As long as the volume of the container  $V$  is considerably larger than atomic dimensions,  $\Omega$  is likely to be considerably larger than  $h^{3N}$ , so the classical description, in terms of phase space, is valid [see the discussion following Eq. (16-6)].

The volume  $\Omega$  can be computed by integrating  $dV_q \, dV_p$  over the region allowed to the system. Since each particle is confined to the volume  $V$ , the integration over the position coordinates is

$$\int \dots \int dV_q = \int \dots \int dx_1 \, dy_1 \, dz_1 \dots dx_N \, dy_N \, dz_N = V^N \quad (17-7)$$

The integration of  $dV_p$  will be discussed in the next chapter; here we shall simply write it as  $\Omega_p$ . Therefore,

$$\Omega = V^N \Omega_p \quad \text{and} \quad S = Nk \ln V + k \ln(\Omega_p / h^3 N) \quad (17-8)$$

Comparison with Eq. (6-5) shows that the entropy of a perfect gas does indeed have a term  $Nk \ln V$  ( $Nk = nR$ ).

Moreover in this case of uniform distribution within  $\Omega$ , the mean energy of the gas,  $U = \sum f_\nu E_\nu$ , will be given by the integral

$$U = \frac{1}{\Omega} \int \dots \int dV_q \int \dots \int H dV_p = \frac{1}{\Omega_p} \int \dots \int H dV_p$$

where

$$H = (1/2m) \sum_{j=1}^N (p_{xj}^2 + p_{yj}^2 + p_{zj}^2)$$

is the total energy of the perfect gas. Thus  $U$  is a function of  $\Omega_p$ , as well as of  $m$  and  $N$  and the shape of the volume in phase space, within which the ensemble is confined; we can emphasize this by writing it as  $U(\Omega_p)$ . Note that this is so only when  $H$  is independent of the  $q$ 's. However, from Eq. (17-8) we have

$$\Omega_p = (h^3/V)^N e^S/k \quad \text{so} \quad U = U(h^3 N V - N e^S/k)$$

Thus the formalism of Eq. (17-4) has enabled us to determine something about the dependence of the internal energy on  $V$  and  $S$  for a system with  $H$  independent of the  $q$ 's. We do not know the exact form of the dependence, but we do know that it is via the product  $V^N e^S/k$ . From this one fact we can derive the equation of state for the perfect gas. We first refer to Eqs. (8-1). If the postulates (17-1) and (17-4) are to correspond to experiment, the partials of the function  $U(h^3 N V - N e^S/k)$  with respect to  $S$  and  $V$  must equal  $T$  and  $-P$ , respectively. But

$$\begin{aligned} \left( \frac{\partial U}{\partial S} \right)_V &= \left( \frac{h^3 N}{k V^N} \right) e^S/k \left( \frac{\partial U}{\partial \Omega_p} \right) \quad \text{and} \\ - \left( \frac{\partial U}{\partial V} \right)_S &= N \left( \frac{h^3 N}{V^{N+1}} \right) e^S/k \frac{\partial U}{\partial \Omega_p} = \frac{k N}{V} \left( \frac{\partial U}{\partial S} \right)_V \end{aligned} \quad (17-9)$$

Thus, if the first partial is to equal the thermodynamic temperature  $T$  and minus the second partial is the pressure  $P$  for any system with  $H$  dependent only on the momenta, the relationship between  $T$ ,  $P$ , and  $V$  must be  $P = (kN/V)T = nRT/V$ , which is the equation of state of a perfect gas. Postulate (17-1) does indeed have ties with "reality."

# 18 The Microcanonical Ensemble

We postponed discussing exactly what volume of momentum space was represented by the quantity  $\Omega_p$  of Eq. (17-8), because we did not need to discuss it at that point. Before we can evaluate  $\Omega_p$  we must take into account the requirements of Liouville's theorem, mentioned in the paragraphs following Eq. (16-3). There it was shown that if the distribution function  $f(q,p)$  had the same value for every point in phase space at which the energy of the system is the same (i.e., over a surface of constant energy), then  $f$  will be independent of time. Thus the finite volume of phase space occupied by the ensemble of Eq. (17-6) must be a region between two constant-energy surfaces, or else just the "area" of a single constant-energy surface.

## Example of a Simple Crystal

The latter alternative, an ensemble of systems, all of which have the same energy, is one that should be investigated further. Since the "volume" of phase space occupied by such an ensemble is finite, the discussion of the previous section applies and for maximum entropy the distribution function should be constant over the entire constant-energy surface. The resulting ensemble, for which  $f(q,p) = 0$  unless  $H(q,p) = U$  (or  $f_\nu = 0$  unless  $E_\nu = U$ , for the case where quantum theory is used), Eq. (17-6) becomes

$$f = 1/\Omega \text{ (or } 1/W) \quad \text{when} \quad H(q,p) = U \text{ (or when } E_\nu = U) \\ \text{zero otherwise} \quad \text{and} \quad S = k \ln(\Omega/h\phi) \text{ (or } k \ln W) \quad (18-1)$$

and is called the *microcanonical ensemble*. Quantity  $\Omega$  is the "area" of the surface in phase space for which  $H(q,p) = U$  and  $W$  is the total number of quantum states that have allowed energies  $E_\nu = U$  (and, when  $\Omega$  is large enough,  $\Omega \simeq h\phi W$ ).

We shall work out the microcanonical distribution for two simple systems. The first is the simple crystal model of Eq. (13-12), with each of the  $N$  atoms in the crystal lattice being a three-dimensional harmonic oscillator with frequency  $\omega/2\pi$ . Here we shall use quantum theory, since the formula for the allowed energies of a quantized oscillator is simple. The allowed energy for the  $i$ -th degree of freedom is  $\hbar\omega(\nu_i + 1/2)$ , where  $\hbar = h/2\pi$  and  $\nu_i$  is the quantum number for the  $i$ -th degree of freedom. Therefore the allowed energy of vibration of this crystal is the sum

$$E_\nu = \hbar\omega \sum_{i=1}^{\phi} \nu_i + \frac{1}{2} \phi \hbar\omega; \quad \phi = 3N \quad (18-2)$$

and the total internal energy, including the potential energy of static compression [see Eq. (13-15)] is

$$U = E_\nu + [(V - V_0)^2/2\kappa V_0] = \hbar\omega M + \frac{3}{2} N \hbar\omega + [(V - V_0)^2/2\kappa V_0] \quad (18-3)$$

where  $M = (\nu_1 + \nu_2 + \dots + \nu_\phi)$  and  $\phi = 3N$ .

A microcanonical ensemble would consist of equal proportions of all the states for which  $M$  is a constant integer, and  $W$  is the number of different permutations of the quantum numbers  $\nu_i$  whose sum is  $M$ . This number can be obtained by induction. When  $\phi = 1$ , there is only one state for which  $\nu_1 = M$ , so  $W = 1$ . When  $\phi = 2$ , there are  $M + 1$  different states for which  $\nu_1 + \nu_2 = M$ , one for  $\nu_1 = 0, \nu_2 = M$ , another for  $\nu_1 = 1, \nu_2 = M - 1$ , and so on to  $\nu_1 = M, \nu_2 = 0$ . When  $\phi = 3$ , there are  $M + 1$  different combinations of  $\nu_2$  and  $\nu_3$  when  $\nu_1$  is 0,  $M$  different ones when  $\nu_1 = 1$ , and so on, so that

$$W = (M + 1) + M + (M - 1) + \dots + 2 + 1 = \frac{1}{2} (M + 1)(M + 2) \quad \text{for } \phi = 3$$

Continuing as before, we soon see that for  $\phi$  different  $\nu$ 's (i.e.,  $\phi$  degrees of freedom),

$$W = \frac{(M + \phi - 1)!}{M!(\phi - 1)!} = \frac{(M + 3N - 1)!}{M!(3N - 1)!} \quad (18-4)$$

From this and from Eq. (18-3) we can obtain  $W$  as a function of  $U$ ;

then from Eq. (18-1) we can obtain  $S$  as a function of  $U$ . Probability  $f_\nu$  that the system has any one of the combinations of the  $\nu$ 's which add up to  $M$  is  $1/W$ .

Since both  $M$  and  $W$  are large integers, we can use the asymptotic formula for the factorial function,

$$n! \simeq (2\pi n)^{1/2} n^n e^{-n}, \quad n \gg 1 \quad (18-5)$$

which is called *Stirling's formula*. Using it, we can obtain a simple approximation for the number  $W$  of different quantum states which have the same value of  $M$ , and thus of  $U$ :

$$\begin{aligned} W &\simeq \left[ \frac{M + \phi - 1}{2\pi M\phi} \right]^{1/2} (M + \phi - 1)^{M + \phi - 1} M^{-M} (\phi - 1)^{-\phi + 1} \\ &\simeq \left[ \frac{\phi}{2\pi M(M + \phi)} \right]^{1/2} \left( 1 + \frac{\phi}{M} \right)^M \left( 1 + \frac{M}{\phi} \right)^\phi \end{aligned} \quad (18-6)$$

where, since  $\phi = 3N$  is large, we have substituted  $\phi$  for  $\phi - 1$ . Therefore the entropy of the simple crystal is

$$S = k \ln W \simeq kM \ln \left( 1 + \frac{\phi}{M} \right) + k\phi \ln \left( 1 + \frac{M}{\phi} \right) \quad (18-7)$$

where we have neglected the logarithm of the square root, since it is so much smaller than the other two terms.

Let us first consider the high-energy case, where the average oscillator is in a higher excited state ( $\nu_i \gg 1$ ) and therefore where  $M \gg \phi = 3N$ . In this case, we can neglect 1 compared to  $M/\phi$  and we can use the approximation  $\ln(1 + x) \simeq x$ , good for  $x \ll 1$ , obtaining

$$S \simeq k\phi + k\phi \ln(M/\phi) = k\phi \ln(eM/\phi)$$

so that

$$(18-8)$$

$$U \simeq 3 \frac{\hbar\omega N}{e} e^{S/3kN} + \frac{3}{2} N\hbar\omega + [(V - V_0)^2 / 2\kappa V_0], \quad \phi = 3N \ll M$$

Remembering that  $T = (\partial U / \partial S)_V$ , we can find  $T$  as a function of  $S$  and then  $U$  as a function of  $T$  and  $V$ , for this high-temperature case,

$$T = \hbar\omega (\partial M / \partial S)_V \simeq (\hbar\omega / ek) e^{S/3kN}$$

so that

$$(18-9)$$

$$U \simeq 3NkT + \frac{3}{2} N\hbar\omega + [(V - V_0)^2 / 2\kappa V_0]$$

This is to be compared with Eq. (13-15). There is a "zero-point" energy  $(3/2)N\hbar\omega$  additional to this formula, but it is small compared to the first term at high temperatures ( $kT \gg \hbar\omega$ ).

At very low energies we have  $\phi \gg M$ , in which case we can write

$$S \approx kM \ln\left(\frac{3eN}{M}\right)$$

so that

$$\frac{1}{T} = \frac{1}{\hbar\omega} \left( \frac{\partial S}{\partial M} \right)_V \approx \frac{k}{\hbar\omega} \ln\left(\frac{\phi}{M}\right) \approx \frac{S}{\hbar\omega M} - \frac{k}{\hbar\omega}$$

or

$$S \approx \frac{\hbar\omega M}{T} + kM = kM \ln\left(\frac{3eN}{M}\right) \quad \text{or} \quad \frac{1}{T} = \frac{k}{\hbar\omega} \ln\left(\frac{3N}{M}\right)$$

or

$$M = 3Ne^{-\hbar\omega/kT}$$

or

$$\begin{aligned} U &\approx 3N\hbar\omega \left( e^{-\hbar\omega/kT} + \frac{1}{2} \right) + [(V - V_0)^2/2\kappa V_0] \\ S &\approx 3Ne^{-\hbar\omega/kT} \left( \frac{\hbar\omega}{T} + k \right) \end{aligned} \tag{18-10}$$

Thus as  $T \rightarrow 0$  ( $M \rightarrow 0$ ) the entropy and the heat capacity  $(\partial U/\partial T)_V$  both go to zero, a major point of difference from the results of Chapter 13. We shall see later [the discussion of Eq. (19-11), for example] that this vanishing of  $S$  and of  $C_V$  as  $T \rightarrow 0$  is a characteristic of systems when classical physics is no longer a good approximation, and the effects of quantization are apparent.

### Microcanonical Ensemble for a Perfect Gas

In the case of a perfect gas of  $N$  point particles in a volume  $V$  of "normal" size, the energy levels are so closely spaced that we can use classical physics for temperatures greater than a fraction of a degree Kelvin. Thus a microcanonical ensemble for such a system is represented by a distribution function  $f(q, p)$  which is zero everywhere in phase space except on the "surface,"

$$H = \sum_{i=1}^{3N} (1/2m)p_i^2 = U \text{ (a constant)} \quad (18-11)$$

where it is  $1/\Omega$ ,  $\Omega$  being the integral of  $dV_Q dV_P$  over this surface (i.e.,  $\Omega$  is the "area" of the surface). The entropy of the micro-canonical ensemble is then given by Eq. (18-1). This classical approximation should be valid for  $T \gg 0.01^\circ K$ , as will be shown later [see the discussion following Eq. (21-5)].

Since the energy of a perfect gas is independent of the positions of the particles, the integral of  $dV_Q$ , as shown in Eq. (17-7), is simply the  $n$ -th power of the volume  $V$  of the container. The integral of  $dV_P$ , however, is the "area" of the surface in momentum space defined by Eq. (18-11). This surface is the  $3N$ -dimensional generalization of a spherical surface; the coordinates are the  $p$ 's and the radius is  $R = (2mU)^{1/2}$ ,

$$p_1^2 + p_2^2 + \dots + p_\phi^2 = R^2; \quad \phi = 3N; \quad R^2 = 2mU$$

Once the area  $\Omega_P$  of this hyperspherical surface is computed, the rest of the calculation is easy, for the volume of phase space occupied is  $\Omega = V^N \Omega_P$  and  $f$  and  $S$  are given by Eq. (17-11).

To find the area we need to define some hyperspherical coordinates. Working by induction:

For two dimensions:

$$x_1 = R \cos \theta_1, x_2 = R \sin \theta_1, x_1^2 + x_2^2 = R^2$$

Element of length of circle:  $ds = R d\theta_1$

For three dimensions:

$$x_1 = R \cos \theta_1, x_2 = R \sin \theta_1 \cos \theta_2, x_3 = R \sin \theta_1 \sin \theta_2$$

Element of area of sphere:  $dA = R^2 \sin \theta_1 d\theta_1 d\theta_2$

For four dimensions:

$$x_1 = R \cos \theta_1, x_2 = R \sin \theta_1 \cos \theta_2, x_3 = R \sin \theta_1 \sin \theta_2 \cos \theta_3,$$

$$x_4 = R \sin \theta_1 \sin \theta_2 \sin \theta_3$$

Element of surface:  $dA = R^3 \sin^2 \theta_1 \sin \theta_2 d\theta_1 d\theta_2 d\theta_3$

For  $\phi$  dimensions:

$$x_1 = R \cos \theta_1, x_2 = R \sin \theta_1 \cos \theta_2, \dots \quad (18-12)$$

$$x_\phi - 1 = R \sin \theta_1 \dots \sin \theta_{\phi-2} \cos \theta_\phi - 1, x_\phi = R \sin \theta_1 \dots \sin \theta_{\phi-2}$$

$$dA = R^{\phi-1} \sin^{\phi-2} \theta_1 \sin^{\phi-3} \theta_2 \dots \sin \theta_{\phi-2} d\theta_1 d\theta_2 \dots d\theta_{\phi-1}$$

where angle  $\theta_{\phi-1}$  goes from 0 to  $2\pi$  and angles  $\theta_1, \dots, \theta_{\phi-2}$  go from 0 to  $\pi$ . To integrate this area element we need the formula

$$\int_0^\pi \sin^n \theta d\theta = \sqrt{\pi} \left[ \left( \frac{1}{2}n - \frac{1}{2} \right)! / \left( \frac{1}{2}n \right)! \right] \quad (18-13)$$

where  $m!$  is the factorial function

$$m! = \int_0^\infty x^m e^{-x} dx = m \cdot (m-1)!; \quad 0! = 1! = 1$$

$$\left( -\frac{1}{2} \right)! = \sqrt{\pi} = 2 \cdot \left( \frac{1}{2} \right)! \quad (18-14)$$

with asymptotic values given by Stirling's formula (18-5). Thus the total area of the hypersphere is [neglecting such factors as  $\sqrt{2}$  and  $\sqrt{1 - (1/\phi)}$ ]

$$\begin{aligned} A = \Omega_p &= R^{\phi-1} \int_0^\pi \sin^{\phi-2} \theta_1 d\theta_1 \dots \int_0^\pi \sin \theta_{\phi-2} d\theta_{\phi-2} \int_0^{2\pi} d\theta_\phi - 1 \\ &= 2\pi R^{\phi-1} \pi^{(\phi-2)/2} \frac{\left( \frac{1}{2}\phi - \frac{3}{2} \right)!}{\left( \frac{1}{2}\phi - 1 \right)!} \frac{\left( \frac{1}{2}\phi - 2 \right)!}{\left( \frac{1}{2}\phi - \frac{3}{2} \right)!} \dots \frac{(0)!}{\left( \frac{1}{2} \right)!} \\ &= \frac{2\pi^{(\phi/2)} R^{\phi-1}}{\left( \frac{1}{2}\phi - 1 \right)!} \\ &\simeq \left( \frac{4\pi m U}{\phi} \right)^{(\phi-1)/2} \left( 1 - \frac{2}{\phi} \right)^{-(\phi/2)} e^{(\phi-2)/2} \\ &\simeq \left( \frac{4\pi m U e}{\phi} \right)^{(\phi/2)}; \quad \phi \gg 1 \end{aligned} \quad (18-15)$$

and the final expression for the volume of phase space occupied is

$$\Omega \simeq V^N (4\pi m U_e / 3N)^{(3/2)N} \quad (18-16)$$

where we have used Eq. (18-5), have replaced  $[(\phi - 1)/2]$  by  $(1/2)\phi = (3/2)N$  and have used the limiting formula for the exponential function

$$\left(1 + \frac{x}{n}\right)^n \rightarrow e^x; \quad n \rightarrow \infty \quad (18-17)$$

and the  $e$  in the formula for  $\Omega$  is, of course, the base of the natural logarithm,  $e = 2.71828$ .

Consequently the entropy of the gas is, from Eq. (17-6),

$$S = Nk \ln [V(4\pi m U_e / 3N h^2)^{3/2}]$$

or

$$U = (3N h^2 / 4\pi m e) V^{-2/3} e^{2S/3Nk} \quad (18-18)$$

which is to be compared with Eq. (17-9) as well as with the discussion following Eq. (6-6). As with the discussion of Eq. (17-9), we can now obtain the thermodynamic temperature and pressure,

$$T = (\partial U / \partial S)_V = (h^2 / 2\pi m e k) V^{-2/3} e^{2S/3Nk} \quad (18-19)$$

$$P = -(\partial U / \partial V)_S = (N h^2 / 2\pi m e) V^{-5/3} e^{2S/3Nk} = NkT/V$$

$$\text{Also, } U = \frac{3}{2} NkT \quad C_V = (\partial U / \partial T)_V = \frac{3}{2} Nk$$

So the microcanonical ensemble does reproduce the thermodynamic behavior of a perfect gas, in complete detail. With Eq. (17-9) we were able to obtain the equation of state, but now that we have computed the dependence of  $\Omega_p$  on  $U$  and  $N$ , the theoretical model also correctly predicts the dependence of  $U$  on  $T$  and thence the heat capacity of the gas.

### The Maxwell Distribution

The microcanonical ensemble can also predict the velocity distribution of molecules in the gas; if we have the distribution function for the whole gas we can obtain from it the distribution function for a constituent particle. Utilizing Eq. (18-19) we can restate our results

thus far. Since  $R = (2mU)^{1/2}$  and  $U = (3/2)NkT$ , we can say that the systems in the microcanonical ensemble for a perfect gas of point particles are uniformly distributed on the surface of a hypersphere in  $3N$ -dimensional momentum space of radius  $(3NmkT)^{1/2}$ . The probability that the point representing the system is within any given region  $dA$  on this surface is equal to the ratio of the area of the region  $dA$  to the total area  $A = \Omega_p$  of the surface. The region near where the  $p_1$  axis cuts the surface, for example, corresponds to the microstates in which the  $x$  component of momentum of particle 1 carries practically all the energy  $(1/2)\phi kT$  of the whole system. It is an interesting property of hyperspheres of large dimensionality (as we shall show) that the areas close to any axis are negligibly small compared to the areas well away from any axis (where the energy is relatively evenly divided between all the degrees of freedom). Therefore the chance that one degree of freedom will turn out to have most of the energy of the whole gas,  $(1/2)\phi kT$ , and that the other components of momentum are zero is negligibly small.

To show this, and incidentally to provide yet another "derivation" of the Maxwell distribution, we note that the probability that the momentum coordinate, which we happen to have labeled by the subscript 1, has a value between  $p_1$  and  $p_1 + dp_1$  can be obtained easily, since we have chosen our angle coordinates such that  $p_1 = R \cos \theta_1$ . Thus the probability

$$dA = \frac{[(1/2)\phi - 1]!}{2\pi(1/2)\phi} \sin\phi - 2 \theta_1 d\theta_1 \sin\phi - 3 \theta_2 d\theta_3 \dots \sin\theta_\phi - 2 d\theta_\phi - 2 d\theta_\phi - 1 \quad (18-20)$$

as a function of  $\theta_1$ , is only large near  $\theta_1 = (1/2)\pi$  [i.e., where  $p_1 = (\phi mkT)^{1/2} \cos \theta_1$  is very small compared to  $(\phi mkT)^{1/2}$ ] and drops off very rapidly, because of the large power of  $\sin \theta_1$ , whenever the magnitude of  $p_1$  increases. The factor  $\sin\phi - 2\theta_1$  ensures that the probability is very small that the degree of freedom labeled 1 carries most of the total kinetic energy  $(1/2)\phi kT$  of the gas. This would be true for each degree of freedom. It is much more likely that each degree of freedom carries an approximately equal share, each having an amount near  $(1/2)kT$ .

The formula for the probability that degree of freedom 1 have momentum between  $p_1$  and  $p_1 + dp_1$ , irrespective of the values of the other momenta, is obtained by integrating  $dA/A$  over  $\theta_2, \theta_3, \dots, \theta_\phi - 1$ . Using the results of Eqs. (18-13) to (18-17) produces

$$f(p_1) dp_1 = \frac{-[(1/2)\phi - 1]!}{\sqrt{\pi} [(1/2)\phi - 3/2]!} \sin \phi - 2 \theta_1 d\theta_1$$

$$= \frac{1}{(\pi \phi mkT)^{1/2}} \frac{[(1/2)\phi - 1]!}{[(1/2)\phi - 3/2]!} \left(1 - \frac{p_1^2}{\phi mkT}\right)^{(\phi - 3)/2} dp_1$$

since  $-d\theta_1 = (1/\phi mkT)^{1/2} (dp_1/\sin \theta_1)$  and  $\sin^2 \theta_1 = 1 - (p_1^2/\phi mkT)$ . To obtain the Maxwell distribution in its usual form we utilize Eqs. (18-5) and (18-17) and consider factors like  $[1 - (2/\phi)]^{1/2}$  and  $[1 - (p_1^2/\phi mkT)]^{-3/2}$  to equal unity [but not such factors to the  $(1/2)\phi$  power, of course]. The calculations go as follows:

$$f(p_1) dp_1 \simeq \frac{1}{(2\pi mkT)^{1/2}} \frac{\left[1 - (2/\phi)\right]^{(\phi/2)}}{\left[1 - (3/\phi)\right]} e^{-p_1^2/2mkT}$$

$$\times [1 - (p_1^2/\phi mkT)]^{(\phi - 3)/2} dp_1$$

$$\simeq \frac{1}{(2\pi mkT)^{1/2}} \exp(-p_1^2/2mkT) dp_1 \quad (18-21)$$

which is the familiar Maxwell distribution for one degree of freedom [see Eq. (12-7)].

This time we arrived at the Maxwell distribution as a consequence of requiring that the total kinetic energy of the gas be  $(1/2)\phi kT$  and that all possible distributions of this energy between the  $\phi$  degrees of freedom be equally likely. For  $\phi = 3N$  large, by far the majority of these configurations represent the energy being divided more or less equally between all  $\phi$  degrees of freedom, with a variance for each  $p$  equal to  $2m$  times the mean kinetic energy per degree of freedom,  $(1/2)kT$ . We note that the Maxwell distribution is not valid unless the individual atoms are, most of the time, unaffected by the other atoms, mutual collisions being rare events.

# 19

## The Canonical Ensemble

The microcanonical ensemble has sufficed to demonstrate that the basic postulates of statistical mechanics correspond to the facts of thermodynamics as well as of kinetic theory. But it has several drawbacks, hindering its general use. In the first place, the computation of the number of microstates that have a given energy is not always easy. It actually would be easier to calculate average values with a distribution function that included a range of energies, rather than one that differs from zero only when the energy has a specific value.

In the second place (and perhaps more importantly) the microcanonical ensemble corresponds to a system with energy  $U$ , completely isolated from the rest of the universe, which is not the way a thermodynamic system is usually prepared. We usually do not know the exact value of the system's energy; we much more often know its temperature, which means that we know its *average* energy. In other words, we do not usually deal with completely isolated systems, but we do often deal with systems kept in contact with a heat reservoir at a given temperature, so that its energy varies somewhat from instant to instant, but its time average is known. This changes the boundary conditions of Eq. (17-2) and the resulting distribution function will differ from that of Eq. (18-1).

### Solving for the Distribution Function

Suppose we prepare an ensemble as follows: Each system has the same number of particles  $N$  and has the same forces acting on the particles. Each system is placed in a furnace and brought to equilibrium at a specified temperature, with each system enclosed in a volume  $V$ . Thus, although we do not know the exact energy of any single system, we do require that the mean energy, averaged over the ensemble, has the relationships between  $S$  and  $T$  expressed in Eqs. (6-3) and (8-8), for example. The distribution function for such

an ensemble, corresponding to a system in contact with a heat reservoir, should satisfy the following requirements:

$$S = -k \sum_{\nu} f_{\nu} \ln f_{\nu} \text{ is maximum,} \quad (19-1)$$

subject to  $\sum_{\nu} f_{\nu} = 1$  and  $\sum_{\nu} f_{\nu} E_{\nu} = U$ ,

the internal energy.

We solve for  $f_{\nu}$  by using Lagrange multipliers. We require

$$S + \alpha_0 \sum_{\nu} f_{\nu} + \alpha_e \sum_{\nu} f_{\nu} E_{\nu} \text{ be maximum,} \quad (19-2)$$

with  $\alpha_0$  and  $\alpha_e$  adjusted so that

$$\sum_{\nu} f_{\nu} = 1 \text{ and } \sum_{\nu} f_{\nu} E_{\nu} = U,$$

where  $U$ , for example, satisfies the equation  $U = F + TS$ . Setting the partials of this function, with respect to the  $f$ 's, equal to zero we obtain

$$-k \ln f_{\nu} - k + \alpha_0 + \alpha_e E_{\nu} = 0$$

or

$$f_{\nu} = \exp[(\alpha_0 - k + \alpha_e E_{\nu})/k] \quad (19-3)$$

The value of the Lagrange multiplier  $\alpha_0$  is adjusted to satisfy the first subsidiary condition,

$$e^{(\alpha_0/k)} - 1 \sum_{\nu} e^{\alpha_0} e^{E_{\nu}/k} = 1; \quad e^{\alpha_0/k} = e/Z; \quad Z = \sum_{\nu} e^{\alpha_0} e^{E_{\nu}/k}$$

The value of the Lagrange multiplier  $\alpha_e$  is obtained by requiring that the sum  $\sum_{\nu} f_{\nu} E_{\nu}$  should behave like the thermodynamic potential  $U$ . For example, the sum we are calling the entropy is related to the sum we are calling  $U$ , by virtue of Eq. (19-3), as follows:

$$S = -k \sum_{\nu} f_{\nu} \ln f_{\nu} = -\sum_{\nu} f_{\nu} (\alpha_0 - k + \alpha_e E_{\nu}) = -(\alpha_0 - k) - \alpha_e U$$

where we have used Eq. (19-3) for  $\ln f_{\nu}$  and also have used the fact that the  $f$ 's must satisfy  $\sum f_{\nu} = 1$  (the first subsidiary condition).

However, if  $U$  is to be the thermodynamic potential of Eqs. (6-3) and (8-8), this relation between  $S$  and  $U$  should correspond to the equation  $S = (-F + U)/T$ . Therefore the Lagrange multipliers must have the following values:

$$\alpha_0 - k = F/T; \quad \alpha_e = -(1/T)$$

The solution of requirements (19-1) is therefore

$$f_\nu = (1/Z)e^{-E_\nu/kT}; \quad Z = \sum_\nu e^{-E_\nu/kT} = e^{-F/kT}$$

$$S = k \sum_\nu f_\nu [\ln Z + (E_\nu/kT)] = \frac{U - F}{T} = -\left(\frac{\partial F}{\partial T}\right)_V \quad (19-4)$$

The ensemble corresponding to this distribution is called the canonical ensemble. The normalizing constant  $Z$ , considered as a function of  $T$  and  $V$ , is called the *partition function*. Part of the computational advantage of the canonical ensemble is the fact that all the thermodynamic functions can be computed from the partition function. For example,

$$F = -kT \ln Z; \quad S = -(\partial F / \partial T)_V; \quad P = -(\partial F / \partial V)_T \quad (19-5)$$

When the separations between successive allowed energies  $E_\nu$  are considerably less than  $kT$ , classical mechanics can be used and instead of sums over the quantum states  $\nu$  of the system we can use integrals over phase space. The distribution function is the probability density  $f(q, p)$ , and, for a system with  $\phi$  degrees of freedom,

$$f(q, p) = (1/h\phi Z) e^{-H(q, p)/kT} \quad (19-6)$$

$$Z = (1/h\phi) \int \dots \int e^{-H/kT} dV_q dV_p$$

where  $H(q, p)$  is the Hamiltonian function of the system, the kinetic plus potential energy, expressed in terms of the  $q$ 's and  $p$ 's [see Eqs. (13-9) and (16-1)]. From  $Z$  one can then obtain  $F$ ,  $S$ , etc., as per Eq. (19-5). The  $H$  of Eq. (19-6) is the total energy of the system, whereas the  $H$  of Eq. (13-10) is the energy of a single molecule. One might say that the canonical distribution function is the Maxwell-Boltzmann distribution for a whole system. It is an exact solution, whereas the  $f$  of Eq. (13-10) for a molecule is only valid in the limit of vanishing interactions between molecules.

### General Properties of the Canonical Ensemble

The first thing to notice about this ensemble is that the distribution function is not constant; all values of the energy are present, but some of them are less likely to occur than others. In general the larger the energy  $H$  the smaller is  $f$ . However, to find out the probability that the system has a particular value of the energy, we must multiply  $f$  by  $\Omega(H)$ , the "area" of the surface of constant  $H$  in phase space. This usually increases rapidly with  $H$ ; for example  $\Omega = Wh^\phi \approx (2\pi eH/3N\omega)^{3N}$  for the simple crystal of Eq. (18-8) and  $\Omega \approx V^N (4\pi mHe/3N)^{(3/2)N}$  for the perfect gas of Eq. (18-16). The product

$\Omega f(q, p) \propto \Omega(H)e^{-H/kT}$  at first increases and then, for  $H$  large enough, the exponential function "takes over" and  $\Omega f$  eventually drops to zero as  $H \rightarrow \infty$ .

The value of  $H$  that has the most representatives in the ensemble is the value for which  $\Omega e^{-H/kT}$  is maximum. For the gas this value is  $H = (3/2)NkT$  and for the crystal it is  $3NkT$ ; in each case it is equal to the average value  $U$  of energy of the ensemble. The number of systems in the ensemble with energy larger or smaller than this mean value  $U$  diminishes quite sharply as  $|H - U|$  increases. Although some systems with  $H \neq U$  do occur, the mean fractional deviation from the mean  $(\Delta H/U)$  of the canonical distribution turns out to be inversely proportional to  $\sqrt{\phi}$  and thus is quite small when  $\phi$  is large. Therefore the canonical ensemble really does not differ very much from the microcanonical ensemble; the chief difference is that it often is easier to handle mathematically.

The advantages of this greater ease are immediately apparent when we wish to determine the general properties of a canonical ensemble. For these can be deduced the properties of the partition function  $Z$ . For example, in many cases the system consists of  $N$  subsystems, the  $i$ -th having  $\delta_i$  degrees of freedom (so that  $\phi = \sum_{i=1}^N \delta_i$ ), each subsystem having negligible interaction with any other, although there may be strong forces holding each subsystem together. For a perfect gas of  $N$  molecules, the molecules are the subsystems, the number of degrees of freedom of each molecule being three times the number of particles per molecule. For a tightly bound crystal lattice the "subsystems" are the different normal modes of vibration of the crystal—and so on. Whenever such a separation is possible, the partition function turns out to be a product of  $N$  factors, one for each subsystem.

To see why this is so, we note that if the subsystems are mutually independent, the Hamiltonian of the system is a sum of  $N$  separate terms,  $H = \sum_{j=1}^N H_j$ , where  $H_j$  is the energy of the  $j$ -th subsystem and is independent of the coordinates of any other subsystem. For a quantized system, the energy  $E_{\nu_1 \nu_2 \dots \nu_\phi}$  is a sum of  $N$  separate terms, the term  $E_j$  being the allowed energy of the  $j$ -th subsystem, dependent on  $\delta_j$  quantum numbers only (call them  $\nu_j, \nu_j + 1, \dots, \nu_j + \delta_j$ ) and independent of all the quantum numbers of the other subsystems. The partition function is the sum of  $\exp(-E_{\nu_1 \nu_2 \dots \nu_\phi}/kT) = e^{-E_1/kT} \dots e^{-E_j/kT} \dots e^{-E_N/kT}$  over all quantum states of all subsystems,

$$Z = \sum_{\text{all } \nu} \exp(-E_{\nu_1 \nu_2 \dots \nu_\phi}/kT) = z_1 \cdot z_2 \dots z_j \dots z_N$$

where

$$z_j = \sum \exp[-E_j(\nu_j, \nu_j + 1, \dots, \nu_j + \delta_j)/kT]$$

(19-7)

the sum for  $z_j$  being over all the quantum numbers of the  $j$ -th subsystem.

For example, the energy of interaction between the magnetic field and the orientation of the atomic magnets in a paramagnetic solid is, to a good approximation, independent of the motion of translation or vibration of these and other atoms in the crystal. Consequently the magnetic term in the Hamiltonian, the corresponding factor in the partition function, and the resulting additive terms in  $F$  and  $S$  can be discussed and calculated separately from all the other factors and terms required to describe the thermodynamic properties of the paramagnetic material. This of course is what was done in Chapter 13 [see Eqs. (13-16) to (13-18)].

### The Effects of Quantization

To follow this general discussion further, we need to say something about the distribution of the quantized energy levels of the  $j$ -th subsystem. There will always be a lowest allowed level, which we can call  $E_{j,1}$ . This may be multiple, of course; there may be  $g_{j1}$  different quantum states, all with this same lowest energy. The next lowest energy can be labeled  $E_{j,2}$ ; it may have multiplicity  $g_{j2}$ ; and so on. Thus we have replaced the set of  $\delta_j$  quantum numbers for the  $j$ -th subsystem by the single index number  $\nu$ , which runs from 1 to  $\infty$ , and for which  $E_{j,\nu+1} > E_{j,\nu}$ , the  $\nu$ -th level having multiplicity  $g_{j\nu}$ .

Thus the  $j$ -th factor in the partition function can be written

$$z_j = \sum_{\nu=1}^{\infty} g_{j\nu} e^{-E_{j\nu}/kT} \quad (19-8)$$

If the energy differences  $E_{j2} - E_{j1}$  and  $E_{j3} - E_{j2}$ , between the lowest three allowed energy levels of the  $j$ -th subsystem are quite large compared to  $kT$ , then the second term in the sum for  $z_j$  is small compared to the first and the third term is appreciably smaller yet, so that

$$z_j \approx g_{j1} e^{-E_{j1}/kT} [1 + (g_{j2}/g_{j1}) e^{-(E_{j2} - E_{j1})/kT}] \quad (19-9)$$

for  $kT$  small compared to  $E_{j2} - E_{j1}$ . The factor in brackets becomes practically independent of  $T$  when  $kT$  is small enough.

The Helmholtz function for the system is a sum of terms, one for each subsystem,

$$F = -kT \ln Z = \sum_{j=1}^N F_j; \quad F_j = -kT \ln z_j \quad (19-10)$$

and the entropy, pressure, and the other thermodynamic potentials are then also sums of terms, one for each subsystem. Whenever any one of the subsystems has energy levels separated farther apart than  $kT$ , the corresponding terms in  $F$ ,  $S$ , and  $U$  have the limiting forms, obtained from Eq. (19-9),

$$F_j \approx -kT \ln g_{j1} + E_{j1} - (g_{j2}/g_{j1})kT e^{-(E_{j2} - E_{j1})/kT}$$

$$S_j \approx k \ln g_{j1} + (g_{j2}/g_{j1}) \left[ k + \frac{E_{j2} - E_{j1}}{T} \right] e^{-(E_{j2} - E_{j1})/kT} \quad (19-11)$$

$$U_j = F_j + S_j T \approx E_{j1} + (g_{j2}/g_{j1})(E_{j2} - E_{j1}) e^{-(E_{j2} - E_{j1})/kT}$$

Thus whenever the  $j$ -th subsystem has a single lowest state ( $g_{j1} = 1$ , i.e., when the subsystem is a simple one) its entropy goes to zero when  $T$  is reduced so that  $kT$  is much smaller than the energy separation between the two lowest quantum states of the subsystem. On the other hand, if the lowest state is multiple,  $S$  goes to  $k \ln g_{j1}$  as  $T \rightarrow 0$ . In either case, however, the heat capacity  $C_{jv} = (\partial U_j / \partial T)_v$  of the subsystem vanishes at  $T = 0$ . Since all the subsystems have non-zero separations between their energy levels, these results apply to all the subsystems, and thus to the whole system, when  $T$  is made small enough. We have thus "explained" the shape of the curve of Fig. 3-1 and the statements made at the beginning of Chapter 9 and the discussion following Eq. (18-10).

### The High-Temperature Limit

When  $T$  is large enough so that many allowed levels of a subsystem are contained in a range of energy equal to  $kT$ , the exponentials in the partition function sum of Eq. (19-8) vary slowly enough with  $\nu$  so that the sum can be changed to a classical integral over phase space, of the form given in Eq. (19-6). In this case, of course, the dependence of  $Z$  on  $T$  and  $V$  is determined by the dependence of the Hamiltonian  $H$  on  $p$  and  $q$ . For example, if the subsystem is a particle in a perfect gas occupying a volume  $V$ ,  $H_j = (p_{jx}^2 + p_{jy}^2 + p_{jz}^2)/2m$  depends only on the momentum, and the factor in the partition function for the  $j$ -th particle is

$$z_j = (1/h^3) \iiint dV_q \iiint \exp [-(p_x^2 + p_y^2 + p_z^2)/2mkT] dV_p$$

$$= (V/h^3)(2\pi mkT)^{3/2} \quad (19-12)$$

and if there are  $N$  particles,

$$F = -NkT \ln V - \frac{3}{2}NkT \ln(2\pi mkT/h^2); \quad U = \frac{3}{2}NkT$$

[but see Eq. (21-13)]

On the other hand, if the "subsystem" is one of the normal modes of vibration of a crystal,  $H_j = (p_j^2/2m) + (m\omega_j^2 q_j^2/2)$  depends on both  $q$  and  $p$ , so that

$$z_j = (1/h) \int e^{-m\omega_j^2 q_j^2/2kT} dq_j \int e^{-p_j^2/2mkT} dp_j \\ = 2\pi kT/\hbar\omega_j = kT/\hbar\omega_j \quad (19-13)$$

and, if there are  $3N$  modes

$$F = kT \sum_{j=1}^{3N} \ln(h\omega_j) - 3NkT \ln(kT); \quad U = 3NkT$$

the difference between  $U = (3/2)NkT$  and  $U = 3NkT$  being caused by the presence of the  $q$ 's in the expression for  $H$  in the latter case.

For intermediate temperatures we may have to use equations like (19-11) for those subsystems with widely spaced levels and classical equations like (19-12) or (19-13) for those with closely packed energy levels. The mean energy of the former subsystems is practically independent of  $T$ , whereas the mean energy of the latter depends linearly on  $T$ : thus only the latter contribute appreciably to the heat capacity of the whole. In a gas of diatomic molecules, for example, the energy levels of translational motion of the molecules are very closely packed, so that for  $T$  larger than  $1^{\circ}\text{K}$ , the classical integrals are valid for the translational motions, but the rotational, vibrational, and electronic motions only contribute to  $C_V$  at higher temperatures.

# 20

## Statistical Mechanics of a Crystal

Two examples of the use of the canonical ensemble will be discussed here; the thermal properties of a crystal lattice and those of a diatomic gas. Both of these systems have been discussed before, but we now have developed the techniques to enable us to work out their properties in detail and to answer the various questions and paradoxes that have been raised earlier.

### Normal Modes of Crystal Vibration

For example, the simplified crystal model of Eq. (13-15) assumed that each atom in a crystal vibrated independently of any other and thus that every atom had the same frequency of vibration. This is obviously a poor approximation for a real crystal, and in this chapter we shall investigate a model in which the effects of one atom's motion on its near neighbors are included, at least approximately. We shall find that this slightly improved model, although still quite simplified, corresponds surprisingly well to the measured thermal behavior of most crystals.

For comparison, however, we shall complete our discussion of the crystal model of Eq. (13-15), with no interaction between atoms and with all atomic frequencies equal. The allowed energy for the  $j$ -th degree of freedom is  $\hbar\omega(\nu_j + 1/2)$  (where  $\nu_j$  is an integer) and the allowed energy of the whole system is

$$E_{\nu_1, \dots, \nu_\phi} = \hbar\omega \sum_{j=1}^{\phi} \nu_j + E_0; \quad E_0 = \frac{1}{2} \phi \hbar\omega + [(V - V_0)^2 / 2\kappa V_0]$$

and therefore the partition function, Helmholtz function, and other quantities are

$$\begin{aligned}
 Z &= e^{-E_0/kT} z_1 z_2 \cdots z_\phi; \quad z_j = \sum_{\nu_j=0}^{\infty} e^{-\hbar\omega\nu_j/kT} = (1 - e^{-\hbar\omega/kT})^{-1} \\
 F &= -kT \ln Z = E_0 + 3NkT \ln(1 - e^{-\hbar\omega/kT}) \\
 S &= -3Nk \ln(1 - e^{-\hbar\omega/kT}) + [(3N\hbar\omega/T)/(e^{\hbar\omega/kT} - 1)] \quad (20-1) \\
 U &= E_0 + [3N\hbar\omega/(e^{\hbar\omega/kT} - 1)] \\
 &\rightarrow \begin{cases} E_0 + 3N\hbar\omega e^{-\hbar\omega/kT} & kT \ll \hbar\omega \\ E_0 + 3NkT & kT \gg \hbar\omega \end{cases} \\
 C_V &= \frac{3N(\hbar\omega)^2 e^{\hbar\omega/kT}}{kT^2 (e^{\hbar\omega/kT} - 1)^2} \rightarrow \begin{cases} (3N\hbar^2\omega^2/kT^2)e^{-\hbar\omega/kT} & kT \ll \hbar\omega \\ 3Nk & kT \gg \hbar\omega \end{cases}
 \end{aligned}$$

Where we have used the formula

$$\sum_{n=0}^{\infty} x^n = 1/(1 - x), \quad |x| < 1 \quad (20-2)$$

to reduce the partition sums to closed formulas. These equations were first obtained by Einstein. The heat capacity  $C_V$  is plotted in Fig. 20-1 as the dashed curve. It does go to zero as  $T$  becomes much smaller than  $\hbar\omega/k$ , as we showed in Chapter 19 that all quantized systems must. Actually it goes to zero more decidedly than the experimental results show actual crystals do. We shall soon see that this discrepancy is caused by our model's neglect of atomic interactions.

In an actual crystal the interaction forces between the atoms, which tend to bring each atom back to its equilibrium position, depend in a complicated way on the displacements of whole groups of atoms. If the displacements are small, the forces depend linearly on the relative displacements and thus the potential energy is a combination of quadratic terms like  $(1/2)K_{ij}q_i^2$ , depending on the displacements from equilibrium of one of the atoms [which were included in the simplified model of Eq. (20-1)] but also terms like  $(1/2)K_{ij}(q_i - q_j)^2$ , corresponding to a force of interaction between one atom and another. Although many of the  $K_{ij}$ 's are small or zero, some are not. The total potential energy is thus

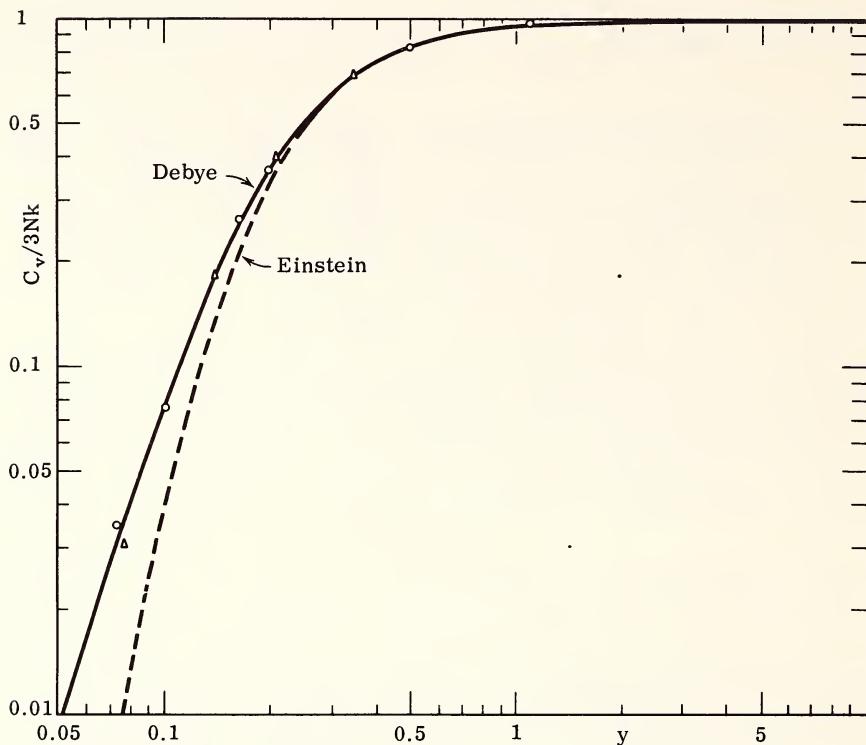


Fig. 20-1. Specific-heat curves for a crystal. Ordinate  $y$  for the Debye curve is  $T/\theta = kT/\hbar\omega_m$ ; ordinate for the Einstein curve is  $3kT/4\hbar\omega$ . Circles are experimental points for graphite, triangles for KCl.

$$\phi = \frac{1}{2} \sum_{i=1}^{3N} \left[ K_i q_i^2 + \sum_{j>i}^{3N} K_{ij} (q_i - q_j)^2 \right] = \frac{1}{2} \sum_{ij=1}^{3N} A_{ij} q_i q_j$$

$$A_{ii} = K_i + 2 \sum_{j \neq i} K_{ij}; \quad A_{ij} = A_{ji} = K_{ij}$$

Therefore the Hamiltonian for the crystal is

$$H = \frac{1}{2m} \sum_{j=1}^{3N} p_j^2 + \frac{1}{2} \sum_{i,j=1}^{3N} A_{ij} q_i q_j \quad (20-3)$$

Actually there are six coordinates not represented in the sum over the  $q$ 's, those for the motion of the crystal as a rigid body; so the total number of coordinates in the second sum is  $3N-6$  rather than  $3N$ . However, 6 is so much smaller than  $3N$  that we can ignore this discrepancy between the sums, by leaving out the kinetic energy of rigid motion and calling  $3N-6$  the same as  $3N$ .

The solution of a dynamical problem of this sort is discussed in all texts of dynamics. The matrix of coefficients  $A_{ij}$  determines a set of normal coordinates,  $Q_n$ , with conjugate momenta  $P_n$ , in terms of which the Hamiltonian becomes a sum of separated terms, each of which is dependent on just one coordinate pair,

$$H = \frac{1}{2} \sum_{n=1}^{3N} [(1/m_n)P_n^2 + m_n\omega_n^2 Q_n^2] + [(V - V_0)^2/2\kappa V_0] \quad (20-4)$$

Application of Hamilton's equations (16-1),  $(\partial H / \partial P_n) = \dot{Q}_n$  and  $(\partial H / \partial Q_n) = -\dot{P}_n$ , results in a set of equations

$$P_n = m_n \dot{Q}_n; \quad \ddot{Q}_n + \omega_n^2 Q_n = 0 \quad (20-5)$$

which may be solved to obtain the classical solution  $Q_n = Q_{0n} e^{i\omega_n t}$ . Thus  $\omega_n/2\pi$  is the frequency of oscillation of the  $n$ -th normal mode of oscillation of the crystal.

These normal modes of the crystal are its various standing waves of free vibration. The lowest frequencies are in the sonic range, corresponding to wavelengths a half or a third or a tenth of the dimensions of the crystal. The highest frequencies are in the infrared and correspond to wavelengths of the size of the interatomic distances. Because there are  $3N$  degrees of freedom there are  $3N$  different standing waves (or rather  $3N-6$  of them, to be pedantically accurate); some of them are compressional waves and some are shear waves.

### Quantum States for the Normal Modes

According to Eq. (16-7), the allowed energies of a single normal mode, with Hamiltonian  $(1/2m_j)P_j^2 + (1/2)m_j\omega_j^2 Q_j^2$  are given by the formula  $\hbar\omega_j[\nu_j + (1/2)]$ , where  $\nu_j$  is an integer, the quantum number of the  $j$ -th normal mode. Sometimes the quantized standing waves are

called *phonons*;  $\nu_j$  is the number of phonons in the  $j$ -th wave. Microstate  $\nu$  of the crystal corresponds to a particular choice of value for each of the  $\nu_j$ 's. The energy of the phonons in microstate  $\nu$  is then

$$E_\nu = E_0(V) + \hbar \sum_{j=1}^{3N} \omega_j \nu_j; \quad E_0 = [(V - V_0)/2\kappa V_0] + \frac{1}{2} \hbar \sum_{j=1}^{3N} \omega_j \quad (20-6)$$

each term in the sum being the energy of a different standing wave. The difference between this and the less accurate Einstein formulas of Eqs. (20-1) is that in the previous case the  $\omega$ 's were the same for all the oscillators, whereas inclusion of atomic interaction in the present model has spread out the resonant frequencies, so that each standing wave has a different value of  $\omega$ .

According to Eq. (19-4) the partition function is

$$Z = \sum_{\text{all } \nu_j \text{'s}} \exp\left(\frac{-E_0 - \hbar \sum_j \omega_j \nu_j}{kT}\right) = e^{-E_0/kT} z_1 z_2 \cdots z_{3N}$$

where

$$z_j = \sum_{\nu_j} e^{-\hbar \omega_j \nu_j / kT} = (1 - e^{-\hbar \omega_j / kT})^{-1} \quad (20-7)$$

and thus, from Eq. (19-5), the Helmholtz function for the crystal is

$$F = -kT \ln Z = E_0(V) + kT \sum_{j=1}^{3N} \ln(1 - e^{-\hbar \omega_j / kT}) \quad (20-8)$$

We can then compute the probability  $f_\nu$  that the system is in the microstate specified by the quantum numbers  $\nu \equiv \nu_1, \nu_2, \dots \nu_{3N}$ . It is the product [see Eq. (19-4)].

$$f_\nu = (1/Z) e^{-E_\nu / kT} = f_1 f_2 f_3 \cdots f_{3N}$$

where

$$f_j = (1/z_j) e^{-\hbar \omega_j \nu_j / kT} = e^{-\hbar \omega_j \nu_j / kT} - e^{-\hbar \omega_j (\nu_j + 1) / kT} \quad (20-9)$$

is the probability that the  $j$ -th standing wave of thermal vibration is

in the  $\nu_j$ -th quantum state. The probability that the crystal is in the microstate  $\nu$  is of course the product of the probabilities that the various normal modes are in their corresponding states.

When  $kT$  is small compared to  $\hbar\omega_j$  for all the standing waves of crystal vibration, all the  $z_j$ 's are practically unity,  $F$  is approximately equal to  $E_0(V)$ , independent of  $T$ , and the entropy is very small. When  $kT$  is large compared to any  $\hbar\omega_j$ , each of the terms in parentheses in Eq. (20-8) will be approximately equal to  $\hbar\omega_j/kT$  and consequently the Helmholtz function will contain a term  $-3NkT \ln(kT)$ , the temperature-dependent term in the entropy will be  $3Nk \ln kT$ , and the heat capacity will be  $3Nk = 3nR$ , as expected. To find values for the intermediate temperatures we must carry out the summation over  $j$  in Eq. (20-8) or, what is satisfactory here, we must approximate the summation by an integral and then carry out the integration.

### Summing over the Normal Modes

The crucial question in changing from sum to integral is: How many standing waves are there with frequencies (times  $2\pi$ ) between  $\omega$  and  $\omega + d\omega$ ? There are three kinds of waves in a crystal, a set of compressional waves and two sets of mutually perpendicular shear waves. If the crystal is a rectangular parallelopiped of dimensions  $l_x, l_y, l_z$ , the pressure distribution of one of the compressional waves would be

$$p = \alpha Q_j \sin(\pi k_j x / l_x) \sin(\pi m_j y / l_y) \sin(\pi n_j z / l_z)$$

where  $Q_j(t)$  is the amplitude of the normal mode  $j$ , with equations of motion (20-5),  $\alpha$  is the proportionality constant relating  $Q_j$  and the pressure amplitude of the compressional wave, and  $k_j, m_j, n_j$  are integers giving the number of standing-wave nodes along the  $x$ ,  $y$ , and  $z$  axes, respectively, for the  $j$ -th wave.

The value of  $\omega_j$ ,  $2\pi$  times the frequency of the  $j$ -th mode, is given by the familiar formula

$$\omega_j^2 = (\pi c k_j / l_x)^2 + (\pi c m_j / l_y)^2 + (\pi c n_j / l_z)^2 \quad (20-10)$$

where  $c$  is the velocity of the compressional wave. Each different  $j$  corresponds to a different trio of integers  $k_j, m_j, n_j$ . A similar discussion will arrive at a similar formula for each of the shear-wave sets, except that the value of  $c$  is that appropriate for shear waves. The problem is to determine how many allowed  $\omega_j$ 's have values between  $\omega$  and  $\omega + d\omega$ .

To visualize the problem, imagine the allowed  $\omega_j$ 's to be plotted as points in "ω space," as shown in Fig. 20-2. They form a lattice

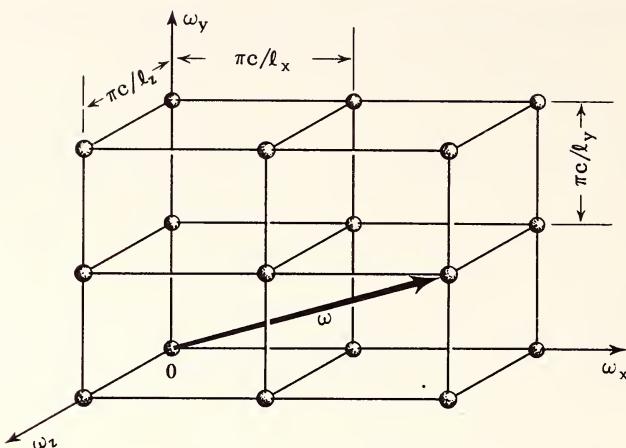


Fig. 20-2. Representation of allowed values of  $\omega$  in  $\omega$  space.

of points in the first octant of the space, with a spacing in the “ $\omega_x$ ” direction of  $\pi c/l_x$ , a spacing in the “ $\omega_y$ ” direction of  $\pi c/l_y$ , and a spacing in the “ $\omega_z$ ” direction of  $\pi c/l_z$ , with the allowed value of  $\omega$  given by the distance from the origin to the point in question, as shown by the form of Eq. (20-10). The point closest to the origin can be labeled  $j = 1$ , the next  $j = 2$ , etc. The spacing between the allowed points is therefore such that there are, on the average,  $l_x l_y l_z / \pi^3 c^3 = V / \pi^3 c^3$  points in a unit volume of “ $\omega$  space,” where  $V = l_x l_y l_z$  is the volume occupied by the crystal.

Therefore all the allowed  $\omega_j$ ’s having value less than  $\omega$  are represented by those points inside a sphere of radius  $\omega$  (with center at the origin). The volume of the part of the sphere in the first octant is  $(1/8)(4\pi\omega^3/3)$  and, because there are  $V/\pi^3 c^3$  allowed points per unit volume, there must be  $(V/\pi^3 c^3)(\pi\omega^3/6)$  standing waves with values of  $\omega_j$  less than  $\omega$ . Differentiating this with respect to  $\omega$ , we see that the average number of  $\omega_j$ ’s with value between  $\omega$  and  $\omega + d\omega$  is

$$d_j = (V/2\pi^2 c^3) \omega^2 d\omega \quad (20-11)$$

Several comments must be made about this formula. In the first place, the formula is for just one of the three sets of standing waves, and thus the  $d_j$  for all the normal modes is the sum of three such formulas, each with its appropriate value of  $c$ , the wave velocity. But we can combine the three by using an average value of  $c$ , and say that, approximately, the total number of standing waves with  $\omega_j$ ’s between  $\omega$  and  $\omega + d\omega$  is

$$dj = (3V/2\pi^2c^3)\omega^2 d\omega \quad (20-12)$$

where  $c$  is an appropriate average of the wave velocities of the compressional and shear waves. Next we should note that Eq. (20-11) was derived for a crystal of rectangular shape. However, a more-detailed analysis of standing waves in crystals of more-general shapes shows that these equations still hold for the other shapes as long as  $V$  is the crystal volume. For a differently shaped crystal, the lattice of allowed points in  $\omega$  space is not that shown in Fig. 20-2, but in spite of this the *density* of allowed points in  $\omega$  space is the same,  $V/\pi^3c^3$ .

Next we remind ourselves that there is an upper limit to the allowed values of the  $\omega_j$ 's; in fact there can only be  $3N$  different normal modes in a crystal with  $N$  atoms ( $3N-6$ , to be pedantically exact). Therefore our integrations should go to an upper limit  $\omega_m$ , where

$$3N = \sum_{j=1}^{3N} 1 = \int_0^{\omega_m} dj = (3V/2\pi^2c^3) \int_0^{\omega_m} \omega^2 d\omega = (V\omega_m^3/2\pi^2c^3)$$

or

$$\omega_m = (6\pi^2 N c^3 / V)^{1/3} \quad (20-13)$$

Finally we note that both Eqs. (20-12) and (20-13) are approximations of the true state of things, first because we have tacitly assumed that  $c$  is independent of  $\omega$ , which is not exactly true at the higher frequencies, and second because we have assumed that the highest compressional frequency is the same as the highest shear frequency, namely,  $\omega_m/2\pi$ , and this is not correct either. All we can do is to hope our approximations tend to average out and that our final result will correspond reasonably well to the measured facts.

### The Debye Formulas

Returning to Eq. (20-8), we change from a sum over  $j$  to an integral over  $dj$ , using Eq. (20-12) and integrating by parts; we obtain

$$\begin{aligned} F &= [(V - V_0)^2 / 2\kappa V_0] + \int_0^{\omega_m} \left[ \frac{1}{2} \hbar \omega_j + kT \ln(1 - e^{-\hbar \omega_j / kT}) \right] dj \\ &= E_0(V) + (3kTV/2\pi^2c^3) \int_0^{\omega_m} \ln(1 - e^{-\hbar \omega / kT}) \omega^2 d\omega \end{aligned}$$

$$= \bar{E}_0 + \frac{kTV\omega_m^3}{2\pi^2 c^3} \ln(1 - e^{-\hbar\omega_m/kT}) - \frac{kTV\omega_m^3}{6\pi^2 c^3} D\left(\frac{\hbar\omega_m}{kT}\right) \quad (20-14)$$

where

$$\bar{E}_0 = [(V - V_0)^2/2\kappa V_0] = \frac{1}{2} \int_0^{\omega_m} \hbar\omega_j dj = (3V\hbar\omega_m^4/16\pi^2 c^3)$$

The function  $D$ , defined by the integral

$$D(x) = (3/x^3) \int_0^x [z^3 dz / (e^z - 1)] \rightarrow \begin{cases} \pi^4/5x^3 & x \gg 1 \\ 1 - (3/8)x & x \ll 1 \end{cases} \quad (20-15)$$

is called the *Debye function*, after the originator of the formula.

We now can express the temperature scale in terms of the *Debye temperature*  $\theta = \hbar\omega_m/k$  (which is a function of  $V$ ) and then write down the thermodynamic functions of interest,

$$F = [(V - V_0)^2/2\kappa V_0] + \frac{9}{8} Nk\theta + NkT[3 \ln(1 - e^{-\theta/T}) - D(\theta/T)]$$

$$\rightarrow \begin{cases} [(V - V_0)^2/2\kappa V_0] + \frac{9}{8} Nk\theta - (\pi^4 NkT^4/5\theta^3) & T \ll \theta \\ [(V - V_0)^2/2\kappa V_0] + \frac{3}{2} Nk\theta + 3NkT \ln(\theta/T) - NkT & T \gg \theta \end{cases}$$

$$S = Nk - 3 \ln(1 - e^{-\theta/T}) + 4D \frac{\theta}{T}$$

$$\rightarrow \begin{cases} (4\pi^4 NkT^3/5\theta^3) & T \ll \theta \\ 3Nk \ln(Te^{4/3}/\theta) & T \gg \theta \end{cases}$$

$$U = [(V - V_0)^2/2\kappa V_0] + \frac{9}{8} Nk\theta + U_V(T); \quad U_V = 3NkT D(\theta/T)$$

(20-16)

$$C_V = 3Nk \left[ 4D \frac{\theta}{T} - \frac{(3\theta/T)}{e^{\theta/T} - 1} \right]$$

$$\rightarrow \begin{cases} (12\pi^4 NkT^3/5\theta^3) & T \ll \theta \\ 3Nk & T \gg \theta \end{cases}$$

$$P = [(V_0 - V)/\kappa V_0] - \frac{9}{8} Nk\theta' - 3NkT(\theta'/\theta)D(\theta/T)$$

$$\xrightarrow{\text{--}} \begin{cases} [(V_0 - V)/\kappa V_0] - \frac{9}{8} Nk\theta' - \frac{3}{5} \pi^4 Nk\theta' (T/\theta)^4 & T \ll \theta \\ [(V_0 - V)/\kappa V_0] - 3NkT(\theta'/\theta) & T \gg \theta \end{cases}$$

where  $\theta' = d\theta/dV = (h/k)(d\omega_m/dV)$  is a negative quantity. Referring to Eq. (3-6) we see that the empirical equation of state is approximately the same as the last line of Eqs. (20-16) if  $-(3Nk\theta'/\theta)$  is equal to  $\beta/\kappa$  of the empirical formula. This relationship can be used to predict values of  $\beta$  if  $\theta'$  can be computed, or it can be used to determine  $\theta'$  from measurements of  $\beta$  and  $\kappa$ .

The functions  $D(x) = [xU_V(\theta/x)/3Nk\theta]$  and  $C_V/3Nk$  are given in Table 20-1 as functions of  $x = \theta/T$ .

Table 20-1

x	D(x)	$C_V/3Nk$	x	D(x)	$C_V/3Nk$
0.0	1.0000	1.0000	4.0	0.1817	0.5031
0.1	0.9627	0.9995	5.0	0.1177	0.3689
0.2	0.9270	0.9980	6.0	0.0776	0.2657
0.5	0.8250	0.9882	8.0	0.0369	0.1382
1.0	0.6745	0.9518	10	0.0193	0.0759
1.5	0.5473	0.8960	12	0.0113	0.0448
2.0	0.4411	0.8259	15	0.0056	0.0230
2.5	0.3540	0.7466	20	0.0024	0.0098
3.0	0.2833	0.6630	25	0.0012	0.0050

A curve of  $C_V/3Nk$  versus  $T/\theta$  is given in Fig. 20-1 (solid curve).

### Comparison with Experiment

Several checks with experiment are possible. By adjusting the value of  $\theta$  we can fit the curve for  $C_V$ , predicted by Eq. (20-16) and drawn in Fig. 20-1, to the experimental curve. That the fit is excellent can be seen from the check between the circles and triangles and the solid line. We see, for example, that the Debye formula, which takes into account (approximately) the coupling between atomic vibrations, fits better than the Einstein formula, which neglects interaction, the discrepancy being greatest at low temperatures.

From the fit one of course obtains an empirical value of  $\theta = \hbar\omega_m/k$  for each crystal measured, and thus a value of  $\omega_m$  for each crystal. However, by actually measuring the standing-wave frequencies of the

crystal and by summing as per Eq. (20-13), we can find out what  $\omega_m$  (and thus  $\theta$ ) ought to be, and then check it against the  $\theta$  that gives the best fit for  $C_V$ . These checks are also quite good, as can be seen from Table 20-2.

Table 20-2

Substance	$\theta, {}^\circ\text{K},$ from $C_V$ fitting	$\theta, {}^\circ\text{K},$ from elastic data
NaCl	308	320
KCl	230	246
Ag	237	216
Zn	308	305

Thus formulas (20-16) represent a very good check with experiment for many crystals. A few differences do occur, however, some of which can be explained by using a somewhat more complicated model. In a few cases, lithium for example, the normal modes are so distributed that the approximation of Eq. (20-12) for the number of normal modes with  $\omega_j$ 's between  $\omega$  and  $\omega + d\omega$  is not good enough, and a better approximation must be used [which modifies Eqs. (20-13) and (20-14)]. In the case of most metals the  $C_V$  does not fit the Debye curve at *very* low temperatures (below about  $2 {}^\circ\text{K}$ ); in this region the  $C_V$  for metals turns out to be more nearly linearly dependent on  $T$  than proportional to  $T^3$ , as the Debye formula predicts. The discrepancy is caused by the free electrons present in metals, as will be shown later.

# 21

## Statistical Mechanics of a Gas

We turn now to the low-density gas phase. A gas, filling volume  $V$ , is composed of  $N$  similar molecules which are far enough apart so the forces between molecules are small compared to the forces within a molecule. At first we assume that the intermolecular forces are negligible. This does not mean that the forces are completely nonexistent; there must be occasional collisions between molecules so that the gas can come to equilibrium. We do assume, however, that the collisions are rare enough so that the mean potential energy of interaction between molecules is negligible compared to the mean kinetic energy of the molecules.

### Factoring the Partition Function

The total energy of the system will therefore be just the sum of the separate energies  $\epsilon(\nu_{\text{mole}})$  of the individual molecules, each one depending only on their own quantum numbers (which we can symbolize by  $\nu_{\text{mole}}$ ) and the partition function can be split into  $N$  molecular factors, as explained in Eq. (19-7):

$$Z = (z_{\text{mole}})^N; \quad z_{\text{mole}} = \sum_{\nu_{\text{mole}}} \exp[-\epsilon(\nu_{\text{mole}})/kT]$$

[but see Eq. (21-12)].

In this case the partition function can be still further factored, for the energy of each molecule can be split into an energy of translation  $H_{\text{tr}}$  of the molecule as a whole, an energy of rotation  $H_{\text{rot}}$ , as a rigid body, an energy of vibration  $H_{\text{vib}}$  of the constituent atoms with respect to the molecular center of mass, and finally an energy of electronic motion  $E_{\text{el}}$ :

$$H_{\text{mole}} = H_{\text{tr}} + H_{\text{rot}} + H_{\text{vib}} + E_{\text{el}} \quad (21-1)$$

To the first approximation these energy terms are independent; the coordinates that describe  $H_{tr}$ , for example, do not enter the functions  $H_{rot}$ ,  $H_{vib}$ , or  $H_{el}$  unless we include the effect of collisions, and we have assumed this effect to be negligible. This independence is not strictly true for the effects of rotation, of course; the rotation does affect the molecular vibration and its electronic states to some extent. But the effects are usually small and can be neglected to begin with.

Consequently, each molecular partition function can be, approximately, split into four separate factors,

$$z_{mole} = z_{tr} \cdot z_{rot} \cdot z_{vib} \cdot z_{el}$$

and the partition function for the system can be divided correspondingly,

$$Z = Z_{tr} \cdot Z_{rot} \cdot Z_{vib} \cdot Z_{el}$$

where

$$Z_{tr} = (z_{tr})^N, \quad z_{rot} = (z_{rot})^N, \quad \text{etc.} \quad (21-2)$$

The individual molecular factors are sums of exponential terms, each corresponding to a possible state of the individual molecule, with quantized energies,

$$z_{tr} = \sum_{k,m,n} \exp(-\epsilon_{kmn}^{tr}/kT), \quad z_{rot} = \sum_{\lambda, \nu} g_{\lambda \nu} \exp(-\epsilon_{\lambda \nu}^{rot}/kT) \quad (21-3)$$

and so on, where  $k, m, n$  are the quantum numbers for the state of translational motion of the molecule,  $\lambda, \nu$  those for rotation, etc., and where  $g_{\lambda \nu}$  are the multiplicities of the rotational states (the  $g$ 's for the translational states are all 1, so they are not written).

The energy separation between successive translational states is very much smaller than the separation between successive rotational states, and these are usually much smaller than the separations between successive vibrational states of the molecule; the separations between electronic states are still another order of magnitude larger. To standardize the formulas, we shall choose the energy origin so the  $\epsilon$  for the first state is zero; thus the first term in each  $z$  sum is unity.

### The Translational Factor

Therefore there is a range of temperature within which several terms in the sum for  $z_{tr}$  are nonnegligible, but only the first term is nonnegligible for  $z_{rot}$ ,  $z_{vib}$ , and  $z_{el}$ . In this range of temperature

the total partition function for the gas system has the simple form

$$Z \approx Z_{tr} = \left[ \sum_{k,m,n} \exp \left( -\epsilon_{kmn}^{tr} / kT \right) \right]^N \quad (21-4)$$

all the other factors being practically equal to unity. To compute  $Z$  for this range of temperature we first compute the energies  $\epsilon_{kmn}$  and then carry out the summation. From it we can calculate  $F$ ,  $S$ , etc., for a gas of low density at low temperatures.

The Schrödinger equation (16-6) for the translational motion of a molecule of mass  $M$  is

$$(\hbar^2/2M)[(\partial^2\Psi/\partial x^2) + (\partial^2\Psi/\partial y^2) + (\partial^2\Psi/\partial z^2)] = -\epsilon^{tr} \Psi$$

If the gas is in a rectangular box of dimensions  $\ell_x, \ell_y, \ell_z$  and volume  $V = \ell_x \ell_y \ell_z$ , with perfectly reflecting walls, the allowed wave functions and energies turn out to be

$$\begin{aligned} \Psi_{kmn} &= A \sin(\pi kx/\ell_x) \cdot \sin(\pi my/\ell_y) \cdot \sin(\pi nz/\ell_z) \\ \epsilon_{kmn}^{tr} &= (\pi^2 \hbar^2/2M)[(k/\ell_x)^2 + (m/\ell_y)^2 + (n/\ell_z)^2] = p^2/2M \end{aligned} \quad (21-5)$$

where  $p$  is the momentum of the molecule in state  $k, m, n$ . For a molecule of molecular weight 30 and for a box 1 cm on a side,  $\pi^2 \hbar^2/2M\ell^2 \approx 10^{-38}$  joule. Since  $k \approx 10^{-23}$  joule/°K, the spacing of the translational levels is very much smaller than  $kT$  even when  $T = 1$  °K, and we can safely change the sum for  $z_{tr}$  into an integral over  $dk$ ,  $dm$ , and  $dn$ ,

$$\begin{aligned} z_{tr} &= \int_0^\infty \int \exp \left\{ -\frac{\hbar^2}{8MkT} \left[ \left( \frac{k}{\ell_x} \right)^2 + \left( \frac{m}{\ell_y} \right)^2 + \left( \frac{n}{\ell_z} \right)^2 \right] \right\} dk dm dn \\ &= (V/h^3)(2\pi MkT)^{3/2} \quad [\text{but see Eq. (21-13)}] \end{aligned} \quad (21-6)$$

by using Eqs. (12-6) (note that we have changed from  $\hbar$  back to  $h = 2\pi\hbar$ ).

This result has been obtained by summing over the quantized states. But with the levels so closely spaced we should not have difficulty in obtaining the same result by integrating over phase space. The translational Hamiltonian is  $p^2/2M$  and the integral is

$$\begin{aligned} z_{tr} &= \int_{-\infty}^{\infty} \int \exp \left[ \frac{-1}{2MkT} (p_x^2 + p_y^2 + p_z^2) \right] h^{-3} dx dy dz dp_x dp_y dp_z \\ &= (V/h^3)(2\pi MkT)^{3/2} \quad [\text{but see Eq. (21-13)}] \end{aligned} \quad (21-7)$$

as before. Integration in (21-7) goes just the same as in (21-6), except that we integrate over  $p_x, p_y, p_z$  from  $-\infty$  to  $+\infty$ , whereas we integrated over  $k, m, n$  from 0 to  $\infty$ ; the result is the same.

The probability  $f_{kmn}$  that a molecule has translational quantum numbers  $k, m, n$  is thus  $(1/z_{tr}) \exp(-\epsilon_{kmn}^{tr}/kT)$  and the probability density that a molecule has translational momentum  $p$  and is located at  $r$  in  $V$  is

$$f(q, p) = (1/V)(2\pi MkT)^{-3/2} \exp(-p^2/2MkT)$$

which is the Maxwell distribution again. Also, in the range of temperature where only  $Z_{tr}$  changes appreciably with temperature, the Helmholtz function and the entropy of the gas are

$$F = -kT \ln(Z_{tr}) = -NkT \left[ \ln(V) + \frac{3}{2} \ln(2\pi MkT/h^2) \right]$$

$$S = Nk \left[ \ln(V) + \frac{3}{2} \ln(2\pi MkT/h^2) \right] + \frac{3}{2} Nk \quad (21-8)$$

[but see Eq. (21-14)]

There is a major defect in this pair of formulas. Neither  $F$  nor  $S$  satisfies the requirement that it be an extensive variable, as illustrated in regard to  $U$  in the discussion preceding Eq. (6-3) (see also the last paragraph in Chapter 8). Keeping intensive variables constant, increasing the amount of material in the system by a factor  $\lambda$  should increase all extensive variables by the same factor  $\lambda$ . If we increase  $N$  to  $\lambda N$  in formulas (21-8), the temperature term will increase by the factor  $\lambda$  but the volume term will become  $\lambda Nk \ln(\lambda V)$ , which is *not*  $\lambda$  times  $Nk \ln(V)$ . The corresponding terms in Eqs. (8-21), giving the thermodynamic properties of a perfect gas, are  $Nk \ln(V/V_0)$ , and when  $N$  changes to  $\lambda N$ ,  $V$  goes to  $\lambda V$  and also  $V_0$  goes to  $\lambda V_0$ , so that the term becomes  $\lambda Nk \ln(\lambda V/\lambda V_0)$ , which is just  $\lambda$  times  $Nk \ln(V/V_0)$ . Evidently the term  $Nk \ln(V)$  in (21-8) should be  $Nk \ln(V/N)$ , or something like it, and thus the partition function of (21-7) should have had an extra factor  $N^{-1}$ , or the partition function for the gas should have had an extra factor  $N^{-N}$  (or something like it). The trouble with the canonical ensemble for a gas seems to be in the way we set up the partition function.

If we remember Stirling's formula (18-5) we might guess that somehow we should have divided the  $Z$  of Eq. (21-1) by  $N!$  to obtain the correct partition function for the gas. The resolution of this dilemma, which is another aspect of Gibbs' paradox, mentioned at the end of Chapter 6, lies in the degree of distinguishability of individual molecules.

### The Indistinguishability of Molecules

Before the advent of quantum mechanics we somehow imagined that, in principle, we could distinguish one molecule from another—that we could paint one blue, for example, so we could always tell which one was the blue one. This is reflected in our counting of translational states of the gas, for we talked as though we could distinguish between the state, where molecule 1 has energy  $\epsilon_1$  and molecule 2 has energy  $\epsilon_2$ , from the state, where molecule 1 has energy  $\epsilon_2$  and molecule 2 has energy  $\epsilon_1$ , for example. But quantum mechanics has taught us that we cannot so distinguish between molecules; a state where molecule 1 has quantum numbers  $k_1, m_1, n_1$ , molecule 2 has  $k_2, m_2, n_2$ , and so on, not only has the same energy as the one where we have reshuffled the quantum numbers among the molecules, it is *really the same state*, and should only be counted once, not  $N!$  times. We have learned that physical reality is represented by the wave function, and that the square of a wave function gives us the probability of presence of a molecule but does not specify *which* molecule is present. Different states correspond to different wave functions, not to different permutations of molecules.

At first sight the answer to this whole set of problems would seem to be to divide  $Z$  by  $N!$ . If particles are distinguishable, there are  $N!$  different ways in which we can assign  $N$  molecules to  $N$  different quantum states. If the molecules are indistinguishable there is only one state instead of  $N!$  ones. This is a good-enough answer for our present purposes. But the correct answer is not so simple as this, as we shall indicate briefly here and investigate in detail later. The difficulty is that, for many states of some systems, the  $N$  particles are not distributed among  $N$  different quantum states; sometimes several molecules occupy the same state.

To illustrate the problem, let us consider a system with five particles, each of which can be in quantum state 0 with zero energy or else in quantum state 1 with energy  $\epsilon$ . The possible energy levels  $E_\nu$  of the system of five particles are, therefore,

$E_0 = 0$	all five particles in lower state
$E_1 = \epsilon$	one particle in upper state, four in lower
$E_2 = 2\epsilon$	two particles in upper state, three in lower
$E_3 = 3\epsilon$	three particles in upper state, two in lower
$E_4 = 4\epsilon$	four particles in upper state, one in lower
$E_5 = 5\epsilon$	all five particles in upper state

(Note that we must distinguish between system states, with energies  $E_\nu$ , and particle states, with energies 0 and  $\epsilon$ .) There is only one system state with energy  $E_0$ , no matter how we count states. All particles are in the lower particle state and there is no question of which particle is in which state. In this respect, a particle state is like the

mathematician's urn, from which he draws balls; ordering of particles inside a single urn has no meaning; they are either in the urn or not.

Distinguishability does come into the counting of the system states having energy  $E_1$ , however. If we can distinguish between particles we shall have to say that five different system states have energy  $E_1$ ; one with particle 1 in the upper state and the others all in the lower "urn," another with particle 2 excited and 1, 3, 4, and 5 in the ground state, and so on. In other words the multiplicity  $g_1$  of Eq. (19-8) is 5 for the system state  $\nu = 1$ . On the other hand, if we cannot distinguish between particles, there is only one state with energy  $E_1$ , the one with one particle excited and four in the lower state (and it has no meaning to ask which particle is excited; they all are at one time or other, but only one is excited at a time).

For distinguishable particles, a count of the different ways we can put five particles into two urns, two in one urn, and three in the other, will show that the appropriate multiplicity for energy  $E_2$  is  $g_2 = 10$ . And so on;  $g_3 = 10$ ,  $g_4 = 5$ ,  $g_5 = 1$ . Therefore, for *distinguishable particles*, the partition function for this simple system would be

$$Z = \sum_{\nu=0}^5 g_{\nu} e^{-\nu \epsilon / kT} = 1 + 5x + 10x^2 + 10x^3 + 5x^4 + x^5$$

$$= (1 + x)^5 \quad \text{where } x = e^{-\epsilon / kT}$$

and where we have used the binomial theorem to take the last step. Thus such a partition function factors into single-particle factors  $z = 1 + e^{-\epsilon / kT}$ , as was assumed in Eqs. (19-8) and (24-3).

On the other hand, if the particles are *indistinguishable*, all the multiplicities  $g$  are unity and

$$Z = 1 + x + x^2 + x^3 + x^4 + x^5$$

which does *not* factor into five single-particle factors.

### Counting the System States

Generalizing, we can say that if we have  $N$  distinguishable particles, distributed among  $M$  different quantum states,  $n_j$  of them in particle state  $j$ , with energy  $\epsilon_j$  (so that  $\sum_{j=1}^M n_j = N$ ), then the number of different ways we can distribute these  $N$  particles among the  $M$  particle states, so that the system energy  $E_{\nu} = \sum_{j=1}^M n_j \epsilon_j$ , is

$$g_\nu = \frac{N!}{n_1! n_2! \cdots n_M!} \quad (21-9)$$

$N!$ , the number of different ways all  $N$  particles can be permuted, being reduced by the numbers  $n_j!$  of different ways the particles could be permuted in each of the  $M$  "urns," since permutation inside an urn does not count. The  $Z$  for distinguishable particles then is

$$\begin{aligned} Z_{\text{dist}} &= \sum_\nu g_\nu \exp \left( - \sum_{j=1}^M n_j \epsilon_j / kT \right) \\ &= \sum_\nu \frac{N!}{n_1! n_2! \cdots n_M!} x_1^{n_1} x_2^{n_2} \cdots x_M^{n_M} = z^N \end{aligned} \quad (21-10)$$

where  $x_j = \exp(-\epsilon_j/kT)$ ,  $z = x_1 + x_2 + \cdots + x_M$ , where the sum is over all values of the  $n_j$ 's for which  $\sum_{j=1}^M n_j = N$ , and where we have used

the multinomial theorem to make the last step. Again this partition function factors into single-particle factors.

Again, if the particles are indistinguishable, the partition function is

$$Z_{\text{ind}} = \sum_\nu x_1^{n_1} x_2^{n_2} \cdots x_M^{n_M} \quad (21-11)$$

with the sum again over all values of the  $n_j$ 's for which  $\sum n_j = N$ . This sum does not factor into single-particle factors.

We thus have reached a basic difficulty with the canonical ensemble. As long as we could consider the particles in the gas as distinguishable, our partition functions came out in a form that could be factored into  $N$   $z$ 's, one for each separate particle. As we have seen, this makes the calculations relatively simple. If we now have to use the canonical ensemble for indistinguishable particles, this factorability is no longer possible, and the calculations become much more difficult. In later chapters we shall find that a more general ensemble enables us to deal with indistinguishable particles nearly as easily as with distinguishable ones. But in this chapter we are investigating whether, under some circumstances, the partition function for the canonical ensemble can be modified so that indistinguishability can approximately be taken into account, still retaining the factorability we have found so useful. Can we divide  $Z_{\text{dist}}$  of Eq. (21-10) by some single factor so it is, at least approximately, equal to the  $Z_{\text{ind}}$  of Eq. (21-11)?

There is a large number of terms in the sum of (21-10) which have multiplicity  $g_\nu = N!$  These are the ones for the system states  $\nu$ , for which all the  $n_j$ 's are 0 or 1, for which no particle state is occupied by more than one particle. We shall call these system states the *sparse* states, since the possible particle states are sparsely occupied. On the other hand, there are other terms in (21-10) with multiplicity less than  $N!$ . These are the terms for which one or more of the  $n_j$ 's are larger than 1; some particle states are occupied by more than one of the particles. Such system states can be called *dense* states, for some particle states are densely occupied. If the number and magnitude of the terms for the sparse states in (21-10) are much larger than the number and magnitude of the terms for the dense states, then it will not be a bad approximation to say that all the  $g_\nu$ 's in (21-10) are equal to  $N!$  and thus that  $(Z_{\text{dist}}/N!)$  does not differ much from the correct  $Z_{\text{ind}}$ . And  $(Z_{\text{dist}}/N!)$  can still be factored, although  $Z_{\text{ind}}$  cannot.

To see when this advantageous situation will occur, we should examine the relative sizes of the terms in the sum of Eq. (21-10). The term for which the factor  $x_1^{n_1} \cdots x_M^{n_M}$  is largest is the one for which  $n_1 = N$ ,  $n_j = 0$  ( $j > 1$ ) (i.e., for which all particles are in the lowest state). This term has the value  $1 \cdot \exp(-N\epsilon_1/kT)$ . It is one of the "densest" states. The largest term for a sparse state is the one for which  $n_1 = n_2 = \cdots = n_N = 1$ ,  $n_j = 0$  ( $j > N$ ) (i.e., for which one particle is in the lowest state, one in the next, and so on up to the  $N$ -th state). Its value is

$$(N!) \exp[-(\epsilon_1 + \epsilon_2 + \cdots + \epsilon_N)/kT]$$

$$\simeq \sqrt{2\pi N} \exp\{N[\ln(N/e) - (\bar{\epsilon}_N/kT)]\}$$

where we have used Stirling's formula (18-5) for  $N!$  and we have written  $\bar{\epsilon}_N$  for the average energy  $[(\epsilon_1 + \epsilon_2 + \cdots + \epsilon_N)/N]$  of the first  $N$  particle states. Consequently, whenever  $\ln(N/e)$  is considerably larger than  $(\bar{\epsilon}_N - \epsilon_1)/kT$ , the sum of sparse-state terms in  $Z_{\text{dist}}$  is so much larger than the sum of dense-state terms that  $Z_{\text{dist}}$  is practically equal to a sum of the sparse-state terms only, and in this case  $Z_{\text{dist}} \simeq N! Z_{\text{ind}}$ . This situation is the case when  $kT$  is considerably larger than the spacing between particle-state energy levels, which is the case when classical mechanics holds.

### The Classical Correction Factor

Therefore whenever the individual particles in the system have energy levels sufficiently closely packed, compared to  $kT$ , so that classical phase-space integrals can be used for at least part of the  $z$

factor, it will be a good approximation to correct for the lack of distinguishability of the molecules by dividing  $Z_{\text{dist}}$  by  $N!$ . In this case there are enough low-lying levels so that each particle can occupy a different quantum state and our initial impulse, to divide  $Z$  by  $N!$ , the number of different ways in which we can assign  $N$  molecules to  $N$  different states, was a good one. Instead of Eq. (21-1) we can use the approximate formula

$$Z \simeq (1/N!)(z_{\text{mole}})^N \simeq (e z_{\text{mole}}/N)^N \quad (21-12)$$

(omitting the factor  $\sqrt{2\pi N}$  in the second form).

Since the translational energy levels of a gas are so closely spaced, this method of correcting for the indistinguishability of the molecules should be valid for  $T > 1^{\circ}\text{K}$ .

The correction factor can be included in the translational factor, so that, instead of Eqs. (21-4) to (21-8), we should use

$$\begin{aligned} Z_{\text{tr}} &= (1/N!)V^N (2\pi MkT/h^2)^{(3/2)N} \simeq (eV/N)^N (2\pi MkT/h^2)^{(3/2)N} \\ &= (eV/n\ell_T^3)^N \end{aligned} \quad (21-13)$$

where  $n = N/N_0$  is the number of moles and where the "thermal length"  $\ell_T = hN_0^{1/3}/\sqrt{2\pi MkT}$  is equal to  $1.47 \times 10^{-2}$  meters for protons at  $T = 1^{\circ}\text{K}$  (for other molecules or temperatures divide by the square root of the molecular weight or of  $T$ ). The values of the translational parts of the various thermodynamic quantities for the gas, corrected for molecular indistinguishability, are then

$$\begin{aligned} F_{\text{tr}} &= -NkT \left[ \ln(eV/N) + \frac{3}{2} \ln(2\pi MkT/h^2) \right] \\ S_{\text{tr}} &= Nk \left[ \ln(V/N) + \frac{3}{2} \ln(2\pi MkT/h^2) \right] + \frac{5}{2} Nk \\ U &= \frac{3}{2} NkT; \quad C_V = \frac{3}{2} Nk; \quad P = (NkT/V) \\ H &= \frac{5}{2} NkT; \quad C_p = \frac{5}{2} Nk \end{aligned} \quad (21-14)$$

The equation for  $S$  is called the *Sackur-Tetrode formula*.

Comparison with Eqs. (8-21) shows that statistical mechanics has indeed predicted the thermodynamic properties of a perfect gas. It has done more, however; it has given the value of the constants of integration  $S_0$ ,  $T_0$ , and  $V_0$  in terms of the atomic constants  $h$ ,  $M$ , and  $k$ , and it has indicated the conditions under which a collection of  $N$  molecules can behave like a perfect gas of point particles.

We also discover that we can now solve Gibbs' paradox, stated at the end of Chapter 6. Mixing two different gases does change the entropy by the amount given in Eq. (6-14). But mixing together two portions of the same gas produces no change in entropy. If the molecules on both sides of the diaphragm are identical, there is really no increase in disorder after the diaphragm is removed. One can never tell (in fact one must never even ask) from which side of the diaphragm a given molecule came, so one cannot say that the two collections of identical molecules "intermixed" after the diaphragm was removed.

We also note that division by  $N!$  was not required for the crystal discussed in Chapter 20. In a manner of speaking,  $N!$  was already divided out. We never tried to include, in our count, the number of ways the  $N$  atoms could be assigned to the different lattice points, and so we did not have to divide out the number again. More will be said about this in Chapter 27.

### The Effects of Molecular Interaction

We have shown several times [see Eqs. (17-9), (18-19), and (21-14)] that, when the interaction between separate molecules in a gas is neglected completely, the resulting equation of state is that of a perfect gas. Before we finish discussing the translational partition function for a gas, we should show how the effects of molecular interaction can be taken into account. We shall confine our discussion to modifications of the translational terms, since these are the most affected. Molecular interactions do change the rotational, vibrational, and electronic motions of each molecule, but the effects are smaller.

The first effect of molecular interactions is to destroy the factorability of the translational partition function, at least partly. The translational energy, instead of being solely dependent on the molecular momenta, now has a potential energy term, dependent on the relative positions of the various molecules. This is a sum of terms, one for each pair of molecules. The force of interaction between molecule  $i$  and molecule  $j$ , to the first approximation, depends only on the distance  $r_{ij}$  between their centers of mass. It is zero when  $r_{ij}$  is large; as the molecules come closer together than their average distance the force is first weakly attractive until, at  $r_{ij}$  equal to twice the "radius"  $r_0$  of each molecule, they "collide" and their closer approach is prevented by a strong repulsive force. Thus the potential energy  $W_{ij}(r_{ij})$  of interaction between molecule  $i$  and molecule  $j$  has the form shown in Fig. 21-1, with a small positive slope (attractive force) for  $r_{ij} > 2r_0$  and a large negative slope (repulsive force) for  $r_{ij} < 2r_0$ . By the time  $r_{ij}$  is as large as the average distance between molecules in the gas,  $W_{ij}$  is zero; in other words we still are assuming that the majority of the time the molecules do not affect each other.

The translational part of the Hamiltonian of the system is

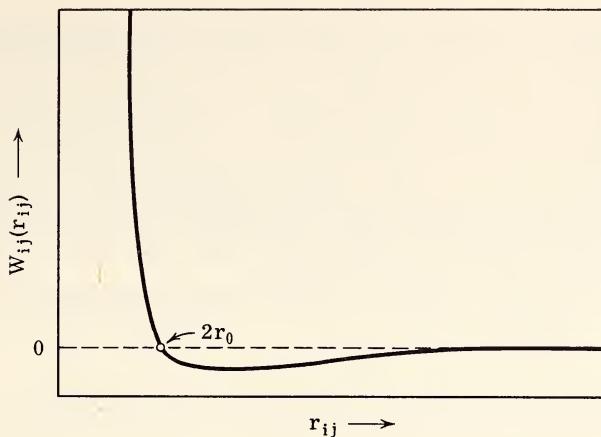


Fig. 21-1. Potential energy of interaction between two molecules as a function of their distance apart.

$$H_{tr} = (1/2M) \sum_{i=1}^N p_i^2 + \sum_{\text{all pairs}} W_{ij}(r_{ij}) \quad (21-15)$$

where the sum of the  $W_{ij}$ 's is over all the  $(1/2)N(N-1) \approx (1/2)N^2$  pairs of molecules in the gas. The translational partition function is then

$$Z_{tr} = \int \cdots \int e^{-H_{tr}/kT} dx_1 dx_2 \cdots dz_N dp_{x1} dp_{y1} \cdots dp_{zN}/h^{3N}$$

The integration over the momentum coordinates can be carried through as with Eq. (21-7) and, since the molecules are indistinguishable, we divide the result by  $N!$ . However the integration over the position coordinates is not just  $V^N$  this time, because of the presence of the  $W_{ij}$ 's,

$$\begin{aligned} Z_{tr} &= Z_p \cdot Z_q; \quad Z_p \approx \frac{1}{N!} \left( \frac{2\pi M k T}{h^2} \right)^{(3/2)N} \\ &\approx \left( \frac{e}{N} \right)^N \left( \frac{2\pi M k T}{h^2} \right)^{(3/2)N} \\ Z_q &= \int \cdots \int \exp \left[ - \sum_{\text{all pairs}} W_{ij}(r_{ij})/kT \right] dx_1 dy_1 \cdots dy_N dz_N \end{aligned} \quad (21-16)$$

Let us look at the behavior of the integrand for  $Z_q$ , as a function of

the coordinates  $x_1, y_1, z_1$  of one molecule. The range of integration is over the volume  $V$  of the container. Over the great majority of this volume the molecule will be far enough away from all other molecules so that  $\sum W_{ij}$  is 0 and the exponential is 1; only when molecule 1 comes close to another molecule (the  $j$ -th one, say) does  $r_{1j}$  become small enough for  $W_{1j}$  to differ appreciably from zero. Of course if  $r_{1j}$  becomes smaller than  $2r_0$ ,  $W_{1j}$  becomes very large positive and the integrand for  $Z_q$  will vanish. The chance that two molecules get closer together than  $2r_0$  is quite small.

Thus it is useful to add and subtract 1 from the integrand,

$$Z_q = \int \cdots \int \{1 + [\exp(-\sum W_{ij}/kT) - 1]\} dx_1 \cdots dz_N \\ = V^N + \int \cdots \int [\exp(-\sum W_{ij}/kT) - 1] dx_1 \cdots dz_N \quad (21-17)$$

where the first unity in the braces can be integrated as in Eq. (21-7) and the second term is a correction to the perfect gas partition function, to take molecular interaction approximately into account. As we have just been showing, over most of the range of the position coordinates the integrand of this correction term is zero. Only when one of the  $r_{ij}$ 's is relatively small is any of the  $W_{ij}$ 's different from zero. To the first approximation, we can assume that only one  $W_{ij}$  differs from zero at a time, as the integration takes place.

Thus the integral becomes a sum of similar integrals, one for each of the  $(1/2)N(N-1) \approx (1/2)N^2$  interaction terms  $W_{ij}$ . A typical one is the integral for which  $W_{1j}$  is not zero; for this one the integrand differs from zero only when  $(x_1, y_1, z_1)$  is near  $(x_j, y_j, z_j)$ , so in the integration over  $dx_1 dy_1 dz_1 = dV_1$  we could use the relative coordinates  $r_{1j}, \theta_{1j}, \phi_{1j}$ . Once this integral is carried out, the integrand for the rest of the integrations is constant, so each of the integrals over the other  $dV_i$ 's is equal to  $V$ . Thus

$$Z_q = V^N + \frac{1}{2} N^2 \int_0^{2\pi} d\phi_{1j} \int_0^\pi \sin \theta_{1j} d\theta_{1j} \int_0^\infty \left( e^{-W_{1j}/kT} - 1 \right) r_{1j}^2 dr_{1j} \\ \times \int \cdots \int dV_2 \cdots dV_N \\ = V^N + \frac{1}{2} N^2 V^{N-1} \left[ 4\pi \int_0^\infty \left( e^{-W_{1j}/kT} - 1 \right) r_{1j}^2 dr_{1j} \right]$$

When  $r_{1j} < 2r_0$ ,  $W_{1j}$  becomes very large positive and the integrand of the last term becomes -1, so this part of the quantity in brackets is just minus the volume of a sphere of radius  $2r_0$ , which we shall call  $-2\beta$ . For  $r_{1j} > 2r_0$ ,  $W_{1j}$  is small and negative, so  $(e^{-W_{1j}/kT} - 1) \approx -(W_{1j}/kT)$ , and this part of the quantity in brackets is roughly

$$-4\pi \int_{2r_0}^{\infty} (W_{1j}/kT) r_{1j}^2 dr_{1j}$$

which we shall call  $2\alpha/kT$ .

### The Van der Waals Equation of State

Therefore, to the first approximation, molecular interaction changes  $Z_{tr}$  from the simple expression of Eq. (21-13) to

$$Z_{tr} \approx \left(\frac{eV}{N}\right)^N \left(\frac{2\pi MkT}{h^2}\right)^{(3/2)N} \left[1 - \left(\frac{N^2\beta}{V}\right) + \left(\frac{N^2\alpha}{VkT}\right)\right] \quad (21-18)$$

where  $N\beta = N(8\pi r_0^3/3)$  is proportional to the total part of the volume  $V$  which is made unavailable to a molecule because of the presence of the other molecules, and where  $\alpha$  is a measure of the attractive potential surrounding each molecule. The  $\beta$  and  $\alpha$  terms in the bracket are both small compared to 1.

The Helmholtz function and the entropy for this partition function are

$$\begin{aligned} F_{tr} &\approx -\frac{3}{2}NkT \ln\left(\frac{2\pi MkT}{h^2}\right) - NkT \ln\left(\frac{eV}{N}\right) - kT \ln\left(1 - \frac{N^2\beta}{V} + \frac{N^2\alpha}{VkT}\right) \\ &\approx -\frac{3}{2}NkT \ln\left(\frac{2\pi MkT}{h^2}\right) - NkT \ln\left(\frac{eV}{N}\right) + NkT\left(\frac{N\beta}{V}\right) - \left(\frac{N^2\alpha}{V}\right) \\ &\approx -\frac{3}{2}NkT \ln\left(\frac{2\pi MkT}{h^2}\right) - NkT \ln\left[\frac{eV}{N}\left(1 - \frac{N\beta}{V}\right)\right] - \left(\frac{N^2\alpha}{V}\right) \\ S_{tr} &\approx \frac{3}{2}Nk + \frac{3}{2}Nk \ln\left(\frac{2\pi MkT}{h^2}\right) + Nk \ln\left[\frac{e}{N}(V - N\beta)\right] \end{aligned} \quad (21-19)$$

Comparison with Eqs. (21-14) shows that  $U$  and  $C_V$  are unchanged, to this approximation, by the introduction of molecular interaction. However the equation of state becomes

$$P = -\left(\frac{\partial F}{\partial V}\right)_T \approx \frac{NkT}{V - N\beta} - \frac{N^2\alpha}{V^2}$$

or

$$\left(P + \frac{N^2\alpha}{V^2}\right)(V - N\beta) \approx NkT = nRT \quad (21-20)$$

which is the Van der Waals equation of state of Eq. (3-4), with  $a = N_0^2\alpha$  and  $b = N_0\beta$ . The correction  $N^2\alpha/V^2$  to  $P$  (which tends to decrease  $P$  for a given  $V$  and  $T$ ) is caused by the small mutual

attractions between molecules; the correction  $N\beta$  to  $V$  (which tends to increase  $P$  for a given  $V$  and  $T$ ) is the volume excluded by the presence of the other molecules. Thus measurement of  $a$  and  $b$  from the empirical equation of state can give us clues to molecular sizes and attractive forces; or else computation of the forces between like molecules can enable us to predict the Van der Waals equation of state that a gas of these molecules should obey.

# 22

## A Gas of Diatomic Molecules

In the molecular gas described in the preceding chapter, as long as  $kT$  is small compared to the energy spacing of rotational quantum levels of individual molecules, only  $Z_{tr}$  differs appreciably from unity and the gas behaves like a perfect gas of point atoms (if we neglect molecular interactions). To see for what temperature range this holds, we need to know the expression for the allowed energies of free rotation of a molecule. This expression is quite complicated for polyatomic molecules, so we shall go into detail only for diatomic molecules.

### The Rotational Factor

If the two constituent nuclei have masses  $M_1$  and  $M_2$  and if they are held a distance  $R_0$  apart at equilibrium, the moment of inertia of the molecule, for rotation about an axis perpendicular to  $R_0$  through the center of mass, is  $I = [M_1 M_2 R_0^2 / (M_1 + M_2)]$ . The moment of inertia about the  $R_0$  axis is zero. The kinetic energy of rotation is then  $1/2I$  times the square of the total angular momentum of the molecule.

This angular momentum is quantized, of course, the allowed values of its square being  $\hbar^2\ell(\ell+1)$ , where  $\ell$  is the rotational quantum number, and the allowed values of the component along some fixed direction in space are one of the  $(2\ell+1)$  values  $-\ell\hbar, -(\ell-1)\hbar, \dots, +(\ell-1)\hbar, +\ell\hbar$ , for each value of  $\ell$ . Put another way, there are  $2\ell+1$  different rotational states which have the energy  $(\hbar^2/2I)\ell(\ell+1)$ , so the partition function for the rotational states of the gas system of  $N$  molecules is

$$Z_{rot} = \left\{ \sum_{\ell=0}^{\infty} (2\ell+1) \exp[-\theta_{rot} \ell(\ell+1)/T] \right\}^N \quad (22-1)$$

where  $\theta_{rot} = \hbar^2/2Ik$ . Therefore when  $T$  is very small compared to  $\theta_{rot}$ ,  $Z_{rot} \approx 1$  and, according to the discussion following Eq. (19-11), the rotational entropy and specific heat are negligible.

Values of  $\theta_{\text{rot}}$  for a few diatomic molecules will indicate at what temperatures  $Z_{\text{rot}}$  begins to be important. For  $\text{H}_2$ ,  $\theta_{\text{rot}} = 85^\circ\text{K}$ ; for  $\text{HD}$ ,  $\theta_{\text{rot}} = 64^\circ\text{K}$ ; for  $\text{D}_2$ ,  $\theta_{\text{rot}} = 47^\circ\text{K}$ ; for  $\text{HCl}$ ,  $\theta_{\text{rot}} = 15^\circ\text{K}$ ; and for  $\text{O}_2$ ,  $\theta_{\text{rot}} = 2^\circ\text{K}$ . Therefore, except for protium (hydrogen), protium deuteride, and deuterium gases,  $T$  is appreciably larger than  $\theta_{\text{rot}}$  in the temperature range where the system is a gas.

In these higher ranges of temperature we can change the sum for  $Z$  into an integral,

$$Z_{\text{rot}} = (z_{\text{rot}})^N$$

$$\simeq \int_0^{\infty} (2\ell + 1) \exp[-\theta_{\text{rot}} (\ell^2 + \ell)/T] d\ell = T/\theta_{\text{rot}}$$

so

$$Z_{\text{rot}} \simeq (T/\theta_{\text{rot}})^N = (8\pi^2 I k T / h^2)^N$$

$$F_{\text{rot}} \simeq -NkT \ln(T/\theta_{\text{rot}}); \quad S_{\text{rot}} \simeq Nk \ln(eT/\theta_{\text{rot}})$$

$$U_{\text{rot}} \simeq NkT; \quad C_v^{\text{rot}} \simeq Nk, \quad T \gg \theta_{\text{rot}} \quad (22-2)$$

Thus for a gas of diatomic molecules at moderate temperatures, where both translational and rotational partition functions have their classical values, the total internal energy is  $(5/2)NkT$  and the total heat capacity is  $(5/2)Nk$ , as mentioned in the discussion following Eq. (13-11). The rotational terms add nothing to the equation of state, however, for the effect of the neighboring molecules on a molecule's rotational states is negligible for a gas of moderate or low densities; consequently  $Z_{\text{rot}}$  and  $F_{\text{rot}}$  are independent of  $V$ . Therefore the equation of state is determined entirely by  $Z_{\text{tr}}$ , unless the gas density is so great that not even the Van der Waals equation of state is valid.

For hydrogen and deuterium, a more careful evaluation of Eq. (22-1) results in

$$Z_{\text{rot}} \rightarrow \begin{cases} \left(1 + 3e^{-2\theta_{\text{rot}}/T}\right)^N, & T < \theta_{\text{rot}} \\ e^{N\theta_{\text{rot}}/4T} \left(\frac{T}{\theta_{\text{rot}}} + \frac{1}{12} + \frac{7\theta_{\text{rot}}}{480T}\right)^N, & T > \theta_{\text{rot}} \end{cases} \quad (22-3)$$

A plot of the exact value of  $C_v^{\text{rot}}/Nk$ , plotted against  $T/\theta_{\text{rot}}$  is shown in Fig. 22-1. We note that  $C_v^{\text{rot}}$  rises somewhat above  $Nk = nR$ , as  $T$  increases, before it settles down to its classical value. The measured

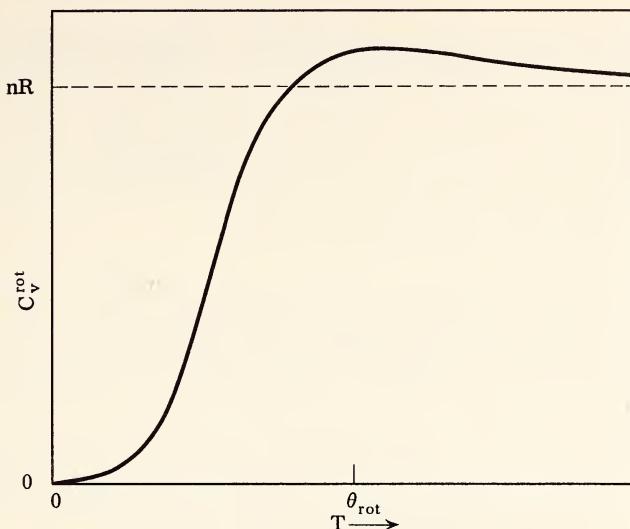


Fig. 22-1. The rotational part of the heat capacity of a diatomic gas as a function of temperature.

values of  $C_v^{\text{rot}}$  for HD fit this curve very well, from  $T = 35^{\circ}\text{K}$  to several hundred degrees K, when molecular vibration begins to make itself felt. But the  $C_v^{\text{rot}}$  curves (i.e.,  $C_v - C_v^{\text{tr}}$ ) for  $\text{H}_2$  and  $\text{D}_2$  do not match, no matter how one juggles the assumed values of  $\theta_{\text{rot}}$ ; for example, the curve for  $\text{H}_2$  has no range of  $T$  for which  $C_v^{\text{rot}} > Nk$ , and the peak for  $\text{D}_2$  is not as large as Fig. 22-1 would predict. The explanation of this anomaly lies again with the effects of indistinguishability of particles. The hydrogen and deuterium homonuclear molecules,  $\text{H}_2$  and  $\text{D}_2$ , are the only ones with a low-enough boiling point so that these effects can be measured. The effects would not be expected for HD, for here the two nuclei in the molecule differ and are thus distinguishable. The calculations for  $\text{H}_2$  and  $\text{D}_2$  will be discussed later, in Chapter 27, after we take up in detail the effects of indistinguishability.

### The Gas at Moderate Temperatures

Therefore, for all gases except  $\text{H}_2$ , HD, and  $\text{D}_2$ , over the temperature range from the boiling point of the gas to the temperature  $\theta_{\text{vib}}$ , where vibrational effects begin to be noticeable, the only effective factors in the partition function are those for translation and rotation, and these factors can be computed classically, using Eqs. (21-14) and (22-2). In this range we can also calculate the partition function for polyatomic molecules. The classical Hamiltonian is  $(p_1^2/2I_1) + (p_2^2/2I_2) + (p_3^2/2I_3)$ , where  $I_1, I_2, I_3$  are the moments of inertia of the molecule about its three principle axes and  $p_1, p_2, p_3$  are the corresponding angular momenta. Therefore,

$$\begin{aligned}
 Z_{\text{rot}} &\simeq \left\{ (8\pi^2/h^3) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp \left( \frac{-(p_1^2/I_1) - (p_2^2/I_2) - (p_3^2/I_3)}{2kT} \right) \right. \\
 &\quad \left. \times dp_1 dp_2 dp_3 \right\}^N \\
 &= [(8\pi^2/\sigma h^3) \sqrt{I_1 I_2 I_3} (2\pi kT)^{3/2}]^N
 \end{aligned} \tag{22-4}$$

where  $8\pi^2$  is the factor produced by the integration over the angles conjugate to  $p_1, p_2, p_3$  and where  $\sigma$  is a symmetry factor, which enters when two or more indistinguishable nuclei are present in a molecule. (If the molecule is asymmetric,  $\sigma = 1$ ; if it has one plane of symmetry,  $\sigma = 2$ ; etc.)

We can now write the thermodynamic functions for a gas for which molecular interactions are negligible, for the temperature range where  $kT$  is large compared with rotational-energy-level differences but small compared with the vibrational-energy spacing;

For monatomic gases, there is no  $Z_{\text{rot}}$  and, from Eq. (21-14),

$$F \simeq F_{\text{tr}} \simeq -NkT \left[ \ln(V/N) + \frac{3}{2} \ln T + F_0 \right]$$

$$U \simeq \frac{3}{2} NkT; \quad C_V \simeq \frac{3}{2} Nk; \quad P \simeq NkT/V$$

For diatomic gases, use Eq. (22-2) for  $Z_{\text{rot}}$ , and

$$F \simeq F_{\text{tr}} + F_{\text{rot}} \simeq -NkT \left[ \ln(V/N) + \frac{5}{2} \ln T + F_0 \right]$$

$$U \simeq \frac{5}{2} NkT; \quad C_V \simeq \frac{5}{2} Nk; \quad P \simeq NkT/V \tag{22-5}$$

For polyatomic gases, use Eq. (22-4) for  $Z_{\text{rot}}$ , and

$$F \simeq F_{\text{tr}} + F_{\text{rot}} \simeq -NkT [\ln(V/N) + 3 \ln T + F_0]$$

$$U \simeq 3NkT; \quad C_V \simeq 3Nk; \quad P \simeq NkT/V$$

where the constant  $F_0$  is a logarithmic function of  $k$ ,  $h$ , the mass  $M$  of the molecule, and of its moments of inertia, the value of which can be computed from Eqs. (21-14) and (22-2) or (22-4). All these formulas are for perfect gases, in that the equation of state is  $PV = NkT$  and the internal energy  $U$  is a function of  $T$  only. The specific heats depend on the nature of the molecule, whether it is monatomic, diatomic, or polyatomic.

We note that the result corresponds to the classical equipartition of energy for translational and rotational motion,  $U$  being  $(1/2)kT$  times the number of "unfrozen" degrees of translational and rotational freedom. The effects of molecular interaction can be allowed for approximately by adding the factor in brackets in Eq. (21-18) to  $Z$ . These results check quite well with the experimental measurements, mentioned following Eq. (13-11).

### The Vibrational Factor

When the temperature is high enough so that  $kT$  begins to equal the spacing between vibrational levels of the molecules, then  $Z_{\text{vib}}$  begins to depend on  $T$  and the vibrational degrees of freedom begin to "thaw out." In diatomic molecules there is just one such degree of freedom, the distance  $R$  between the nuclei. The corresponding potential energy  $W(R)$  has its minimum value at  $R_0$ , the equilibrium separation between the nuclei, and has a shape roughly like that shown in Fig. 22-2. As  $R \rightarrow \infty$ , the molecule dissociates into separate atoms;

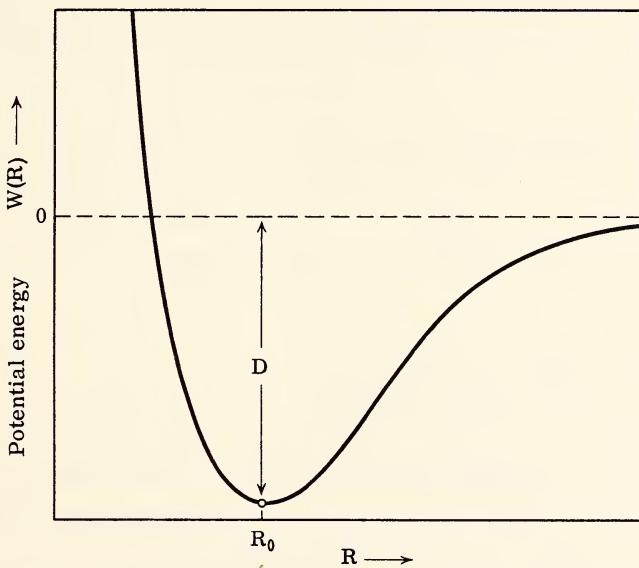


Fig. 22-2. Diatomic molecular energy  $W(R)$  as a function of the separation  $R$  between nuclei.

the energy required to dissociate a molecule from equilibrium is  $D$ , the dissociation energy.

If the molecule is rotating there will be added a dynamic potential, corresponding to the centrifugal force, which is proportional to the square of the molecule's angular momentum and inversely proportional

to  $R^3$ . Fortunately, for most diatomic molecules, this term, which would couple  $Z_{\text{rot}}$  and  $Z_{\text{vib}}$ , is small enough so we can neglect it here. For small-amplitude vibrations about  $R_0$  the system acts like a harmonic oscillator, with a natural frequency  $\omega/2\pi$  which is a function of the nuclear masses and of the curvature of the  $W(R)$  curve near  $R_0$ . Thus the lower energy levels are  $\hbar\omega(n + 1/2)$ , where  $n$  is the vibrational quantum number.

Therefore, to the degree of approximation which neglects coupling between rotation and vibration and which considers all the vibrational levels to be those of a harmonic oscillator,

$$Z_{\text{vib}} \simeq \left( e^{-\epsilon_0/kT} \sum_{n=0}^{\infty} e^{-n\hbar\omega/kT} \right)^N = \left[ e^{+\epsilon_0/kT} \left( 1 - e^{-\hbar\omega/kT} \right) \right]^{-N} \quad (22-6)$$

where  $\epsilon_0 = W(R_0) + (1/2)\hbar\omega$ , and where we have used Eq. (20-2) to reduce the sum. The corresponding contributions to the Helmholtz function, entropy, etc., of the gas are

$$F_{\text{vib}} \simeq N\epsilon_0 + NkT \ln \left( 1 - e^{-\hbar\omega/kT} \right)$$

$$U_{\text{vib}} \simeq N\epsilon_0 + \frac{N\hbar\omega}{e^{+\hbar\omega/kT} - 1} \rightarrow \begin{cases} N\epsilon_0 + N\hbar\omega e^{-\hbar\omega/kT} & kT \ll \hbar\omega \\ N\epsilon_0 + NkT & kT \gg \hbar\omega \end{cases}$$

$$C_v^{\text{vib}} \simeq \frac{N\hbar^2\omega^2}{kT^2} \frac{e^{\hbar\omega/kT}}{\left( e^{\hbar\omega/kT} - 1 \right)^2} \rightarrow \begin{cases} (N\hbar^2\omega^2/kT^2) e^{-\hbar\omega/kT} & kT \ll \hbar\omega \\ Nk & kT \gg \hbar\omega \end{cases} \quad (22-7)$$

which are added to the functions of Eqs. (22-5) whenever the temperature is high enough (for  $T$  equal to or larger than  $\theta_{\text{vib}} = \hbar\omega/k$ ).

As examples of the limits above which these terms become appreciable, the quantity  $\theta_{\text{vib}}$  is equal to  $2200^{\circ}\text{K}$  for  $\text{O}_2$ , to  $4100^{\circ}\text{K}$  for  $\text{HCl}$  and to  $6100^{\circ}\text{K}$  for  $\text{H}_2$  (Fig. 22-3). Therefore below roughly  $1000^{\circ}\text{K}$  the contribution of molecular vibration to  $S$ ,  $U$ , and  $C_v$  of diatomic gases is small. Above several thousand degrees, the vibrational degree of freedom becomes "unfrozen," an additional energy  $kT$  is added per molecule, and an additional  $Nk$  to  $C_v$  [a degree of

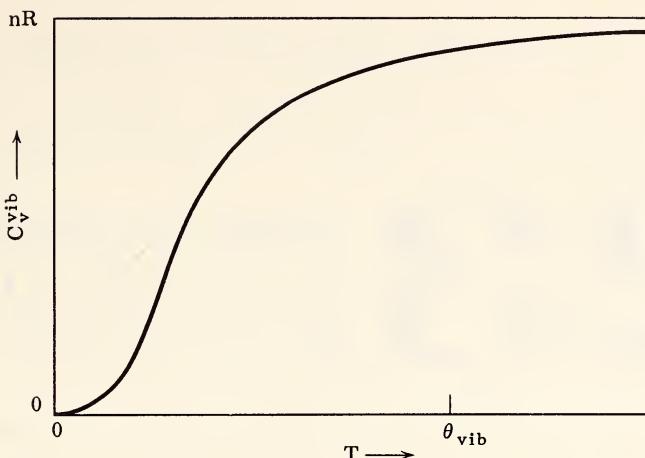


Fig. 22-3. Vibrational part of the heat capacity of a diatomic molecular gas as a function of temperature.

freedom with quadratic potential classically has energy  $kT$ ; see Eq. (13-14)].

In the case of a polyatomic molecule with  $n$  nuclei, there are  $3n - 6$  vibrational degrees of freedom, each with its fundamental frequency  $\omega_j/2\pi$ . The vibrational partition function is

$$Z_{\text{vib}} = e^{-N\epsilon_0} Z_1 Z_2 \cdots Z_{3n-6} \quad \text{where}$$

$$Z_j \simeq \left(1 - e^{-\hbar\omega_j/kT}\right)^{-N} \quad (22-8)$$

[compare this with Eq. (20-7)]. Again, for polyatomic gases, the vibrational contribution below about  $1000^\circ\text{K}$  is small, at higher  $T$  the contribution to  $U$  is  $N(3n - 6)kT$ . It often happens that the molecules are dissociated into their constituent atoms before the temperature is high enough for the vibrational term to "unfreeze."

The temperature would have to be still higher before  $Z_{\text{el}}$  began to have any effect. The usual electronic-level separation divided by  $k$  is roughly equal to  $10,000^\circ\text{K}$  at which temperatures most gases are dissociated and partly ionized. Such cases are important in the study of stellar interiors, but are too complex to discuss in this book. And before we can discuss the thermal properties of electrons we must return to first principles again.

# 23

## The Grand Canonical Ensemble

The canonical ensemble, representing a system of  $N$  particles kept at constant temperature  $T$ , has proved to be a useful model for such systems as the simple crystal and the perfect (or nearly perfect) gas. Many other systems, more complicated than these, can also be represented by the canonical ensemble, which makes it possible to express their thermodynamic properties in terms of their atomic structure. But in Chapter 21 we discovered a major defect, not in the accuracy of the canonical ensemble when correctly applied, but in its ease of manipulation in some important cases.

Whenever the  $N$  particles making up the system are identical and indistinguishable, the corresponding change in the multiplicity factors  $g_V$  has the result that the correct partition function does not separate into a product of  $N$  independent factors, even if the interaction between particles is negligible. In cases where  $kT$  is large compared to the separation between quantum levels of the system, we found we could take this effect into account approximately by dividing by  $N!$ . In this chapter we shall discuss a more general kind of ensemble, which will allow us to retain factorability of partition function and at the same time take indistinguishability into account exactly, no matter what value  $T$  has.

### An Ensemble with Variable $N$

The new ensemble, which we shall call the *grand canonical ensemble*, is one in which we relax the requirement that we placed on the microcanonical and canonical ensembles—that each system in the ensemble has exactly  $N$  particles. We can imagine an infinitely large, homogeneous supersystem kept at constant  $T$  and  $P$ . The system the new ensemble will represent is that part of the supersystem contained within a volume  $V$ . We can imagine obtaining one of the sample systems of the ensemble by withdrawing that part of the supersystem which happens to be in a volume  $V$  at the instant of removal, and of

doing this successively to obtain all the samples that make up the ensemble. Not only will each of the samples differ somewhat in regard to their total energy, but the number of particles  $N$  in each sample will differ from sample to sample. Only the average energy  $U$  and the average number of particles  $\bar{N}$ , averaged over the ensemble, will be specified.

The equations and subsidiary conditions serving to determine the distribution function are thus still more relaxed than for the canonical ensemble. A microstate of the grand canonical ensemble is specified by the number of particles  $N$  that the sample system has, and by the quantum numbers  $\nu_N \equiv \nu_1, \nu_2, \dots, \nu_{3N}$ , which the sample may have and which will specify its energy  $E_{N\nu}$ . Thus for an equilibrium macrostate the distribution function  $f_{N\nu}$  must satisfy the following requirements:

$$S = -k \sum_{N,\nu} f_{N\nu} \ln f_{N\nu} \text{ is maximum,} \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} (23-1)$$

subject to

$$\sum_{N,\nu} f_{N\nu} = 1; \quad \sum_{N,\nu} E_{N\nu} f_{N\nu} = U; \quad \sum_{N,\nu} N f_{N\nu} = \bar{N}$$

where  $U$ ,  $\bar{N}$ , and  $S$  are related by the usual thermodynamic relationships, such as  $U = TS + \Omega + \bar{N}\mu$ , for example, or any other of the equations (8-21). Function  $\Omega$  is the grand potential of Eq. (8-15). Note that, instead of  $n$ , the mean number of moles in the system, we now use  $\bar{N}$ , the mean number of particles, and therefore  $\mu$  is now the chemical potential (the Gibbs function) *per particle*, rather than per mole, as it was in the first third of this book. We shall consistently use it thus henceforth, so it should not be confusing to use the same symbol,  $\mu$ .

### The Grand Partition Function

As before, we simplify the requirements by using Lagrange multipliers, and require that

$$-k \sum_{N,\nu} f_{N\nu} \ln f_{N\nu} + \alpha_1 \sum_{N,\nu} f_{N\nu} + \alpha_e \sum_{N,\nu} E_{N\nu} f_{N\nu} + \alpha_n \sum_{N,\nu} N f_{N\nu} \text{ be maximum,} \quad (23-2)$$

with  $\alpha_1, \alpha_e, \alpha_n$  chosen so that

$$\sum_{N,\nu} f_{N\nu} = 1; \quad \sum_{N,\nu} E_{N\nu} f_{N\nu} = U; \quad \sum_{N,\nu} N f_{N\nu} = \bar{N} \quad (23-2)$$

The partials with respect to the  $f_{N\nu}$ 's, which must be made zero, result in the equations

$$k \ln f_{N\nu} + k = \alpha_1 + \alpha_n N + \alpha_e E_{N\nu}$$

or

$$f_{N\nu} = \exp [(1/k)(\alpha_1 - k + \alpha_n N + \alpha_n E_{N\nu})] \quad (23-3)$$

The first requirement is met by setting

$$e^{(k - \alpha_1)/k} = \mathcal{Q} = \sum_{N,\nu} e^{(\alpha_e N + \alpha_e E_{N\nu})/k}$$

The other two are met by inserting this into the expression for  $S$ ,

$$S = - \sum_{N,\nu} f_{N\nu} (\alpha_1 - k + \alpha_n N + \alpha_e E_{N\nu}) = (k - \alpha_1) - \alpha_n \bar{N} - \alpha_e U$$

and then identifying this with the equations  $S = (U - \Omega - \bar{N}\mu)/T$ , from Eq. (8-21).

We see that we must have

$$k - \alpha_1 = -(\Omega/T) = k \ln \mathcal{Q}; \quad \alpha_n = \mu/T; \quad \alpha_e = -(1/T)$$

so that the solution of Eq. (23-1) is

$$f_{N\nu} = (1/\mathcal{Q}) \exp \left( \frac{\mu N - E_{N\nu}}{kT} \right); \quad \mathcal{Q} = \sum_{N,\nu} \exp \left( \frac{\mu N - E_{N\nu}}{kT} \right)$$

$$\Omega = -kT \ln \mathcal{Q} = -PV; \quad (\partial \Omega / \partial \mu)_{TV} = -\bar{N}$$

$$(\partial \Omega / \partial T)_{V\mu} = -S; \quad (\partial \Omega / \partial V)_{T\mu} = -P \quad (23-4)$$

$$C_V = T(\partial S / \partial T)_{V\mu}; \quad F = \Omega + \mu \bar{N}; \quad U = F + ST = \Omega + ST + \mu \bar{N}$$

These are the equations for the grand canonical ensemble. The sum  $\mathcal{Q}$  is called the grand partition function; it is the sum of the canonical partition functions  $Z(N)$  for ensembles with different  $N$ 's, with weighting factors  $e^{\mu N / kT}$ ,

$$\mathcal{Z} = \sum_{N=0}^{\infty} e^{\mu N/kT} Z(N); \quad Z(N) = \sum_{\nu} e^{-E_{N\nu}/kT} \quad (23-5)$$

All the thermodynamic properties of the system can be obtained from  $\Omega$  by differentiation, as with the canonical ensemble. We shall see that this partition function has even greater possibilities for factoring than does its canonical counterpart.

### The Perfect Gas Once More

Just to show how this ensemble works we take up again the familiar theme of the perfect gas of point particles. From Eq. (21-13) we see that, if we take particle indistinguishability approximately into account, the canonical partition function for the gas of  $N$  particles is

$$Z(N) \simeq (1/N!) \left( V/\ell_t^3 \right)^N; \quad \ell_t = (h/\sqrt{2\pi M kT})$$

and therefore the grand partition function is, from Eq. (23-5),

$$Z = \sum_{N=0}^{\infty} (1/N!) \left[ (V/\ell_t^3) e^{\mu/kT} \right]^N = \exp \left[ (V/\ell_t^3) e^{\mu/kT} \right] \quad (23-6)$$

where we have used the series expression  $e^x = \sum_{n=0}^{\infty} (x^n/n!)$ .  
Then, from Eqs. (23-4)

$$\Omega = -kTV(2\pi MkT/h^2)^{3/2} e^{\mu/kT} = -PV$$

$$\bar{N} = V(2\pi MkT/h^2)^{3/2} e^{\mu/kT} = -(\Omega/kT) = PV/kT$$

$$S = kV(2\pi MkT/h^2)^{3/2} e^{\mu/kT} \left( \frac{5}{2} - \frac{\mu}{kT} \right) = \bar{N}k \left( \frac{5}{2} - \frac{\mu}{kT} \right)$$

$$U = \Omega + ST + \mu \bar{N} = -\bar{N}kT + \bar{N}kT \left( \frac{5}{2} - \frac{\mu}{kT} \right) + \mu \bar{N} = \frac{3}{2} \bar{N}kT$$

$$F = \bar{N}(\mu - kT); \quad \mu = -kT \ln [(V/\bar{N})(2\pi MkT/h^2)^{3/2}] \quad (23-7)$$

which of course present, in slightly different form, the same expressions for  $U$ ,  $C_V$ , and the equation of state as did the other ensembles; only now  $\bar{N}$  occurs instead of  $N$ . We also obtain directly an expression for the chemical potential per particle,  $\mu$ , for the perfect gas.

The probability density that a volume  $V$  of such a gas, in equilibrium at temperature  $T$  and chemical potential  $\mu$ , happens to contain

$N$  particles, and that these particles should have momenta  $p_1, p_2, \dots, p_N$  and be located at the points specified by the vectors  $r_1, r_2, \dots, r_N$ , is then

$$f_N(q, p) = \left( \frac{h^{-3N}}{N!} \right) \exp \left\{ \frac{1}{kT} \left[ \mu N - \sum_{j=1}^N \left( \frac{p_j^2}{2M} \right) \right] - V \left( \frac{2\pi MkT}{h^2} \right)^{3/2} e^{\mu/kT} \right\} \quad (23-8)$$

This is a generalization of the Maxwell distribution. The expression not only gives us the distribution in momentum of the  $N$  particles which happen to be in volume  $V$  at that instant (it is of course independent of their positions in  $V$ ), but it also predicts the probability that there will be  $N$  molecules in volume  $V$  then. If we should wish to use  $P$  and  $T$  to specify the equilibrium state, instead of  $\mu$  and  $T$ , this probability density would become

$$f_N(q, p) = \frac{1}{N!} \frac{(P/kT)^N}{(2\pi MkT)^{(3/2)N}} \exp \left( - \sum_{j=1}^N \frac{p_j^2}{2MkT} - \frac{PV}{kT} \right) \quad (23-9)$$

### Density Fluctuations in a Gas

By summing  $f_{N\nu}$  over  $\nu$  for a given  $N$  (or by integrating  $f_N(q, p)$  over the  $q$ 's and  $p$ 's for a given  $N$ ) we shall obtain the probability that a volume  $V$  of the gas, at equilibrium at pressure  $P$  and temperature  $T$ , will happen to have  $N$  molecules in it. From Eqs. (23-4) and (23-5) this is

$$f_N = \sum_{\nu} f_{N\nu} = (1/2) e^{\mu N/kT} Z(N) = e^{(\Omega + \mu N)/kT} Z(N)$$

Using the expressions for  $Z(N)$  and those for  $(\Omega/kT)$  and  $(\mu/kT)$ , we obtain

$$f_N = e^{-\bar{N}} \left[ \left( \frac{\bar{N}}{V} \right) \left( \frac{h^2}{2\pi MkT} \right)^{3/2} \right]^N \left( \frac{1}{N!} \right) \left[ V \left( \frac{2\pi MkT}{h^2} \right)^{3/2} \right]^N$$

$$= (\bar{N}^N / N!) e^{-\bar{N}} = (1/N!) (PV/kT)^N e^{-PV/kT} \quad (23-10)$$

This is a Poisson distribution [see Eq. (11-5)] for the number of particles in a volume  $V$  of the gas. The mean number of particles is  $\bar{N} = PV/kT$  and the probability  $f_N$  is greatest for  $N$  near  $\bar{N}$  in value. But  $f_N$  is not zero when  $N$  differs from  $\bar{N}$ ; it is perfectly possible to find a volume  $V$  in the gas which has a greater or smaller number of molecules in it than  $PV/kT$ . The variance of the number present is

$$(\Delta N)^2 = \sum_{N=0}^{\infty} (N - \bar{N})^2 f_N = \sum_{N=0}^{\infty} N^2 f_N - (\bar{N})^2 = \bar{N} = PV/kT \quad (23-11)$$

and the fractional deviation from the mean is

$$\Delta N/\bar{N} = \sqrt{\bar{N}} = 1/\sqrt{kT/PV} \quad (23-12)$$

(It should be remembered that the system described by the grand canonical ensemble is not a gas of  $N$  molecules in a volume  $V$ , but that part of a supersystem which happens to be in a volume  $V$ , where  $V$  is much smaller than the volume occupied by the supersystem; thus the number of particles  $N$  that might be present can vary from zero to practically infinity.)

The smaller the volume of the gas looked at (the smaller the value of  $\bar{N}$ ) the greater is this fractional fluctuation of number of particles present (or of density, for  $\Delta N/\bar{N} = \Delta\rho/\rho$ ). Thus we have arrived at the result of Eq. (15-6), for the density fluctuations in various portions of a gas, by a quite different route.

# 24

## Quantum Statistics

But we still have not demonstrated the full utility of the grand canonical ensemble for handling calculations involving indistinguishable particles. The example in the previous chapter used the approximate correction factor  $(1/N!)$ , which we saw in Chapter 21 was not valid at low temperatures or high densities. We must now complete the discussion of the counting of states, which was begun there.

### Occupation Numbers

In comparing the partition functions for distinguishable and indistinguishable particles, given in Eqs. (21-10) and (21-11) for the canonical ensemble, we saw that it was easier to compare the two if we talked about the number of particles occupying a given particle state rather than talking about which particle is in which state. In fact if the particles are indistinguishable it makes no sense to talk about which particle is in which state. We were there forced to describe the system state  $\nu$  by specifying the number  $n_j$  of particles which occupy the  $j$ -th particle state, each of them having energy  $\epsilon_j$ . The number  $n_j$  are called *occupation numbers*.

Of course if the interaction between particles is strong (as is the case with a crystal) we cannot talk about separate particle states; occupation numbers lose their specific meaning and we have to talk about normal modes instead of particles. But let us start with the particle interactions being small enough so we can talk about particle states and their occupation numbers. The results we obtain will turn out to be capable of extension to the strong-interaction case.

We thus assume that, in the system of  $N$  particles, it makes sense to talk about the various quantum states of an individual particle, which we call particle states. These states are ranked in order of increasing energy, so that if  $\epsilon_j$  is the energy of a particle in state  $j$ , then  $\epsilon_{j+1} > \epsilon_j$ . Instead of specifying the system state  $\nu$  by listing what state particle 1 is in, and so on for each particle, we specify it by saying

how many particles are in state  $j$  (i.e., by specifying  $n_j$ ). Thus when the system is in state  $\nu \equiv (n_1, n_2, \dots, n_j, \dots)$ , the total number of particles and the total energy are

$$N = \sum_j n_j; \quad E_N = \sum_j n_j \epsilon_j \quad (24-1)$$

For the canonical ensemble, we have to construct the partition function  $Z$  for a system with exactly  $N$  particles; the sum over  $\nu$  includes only those values of the  $n_j$ 's for which their sum comes out to equal  $N$ . This restriction makes the calculation of a partition function like that of Eq. (21-11) more difficult than it needs to be. With the grand canonical ensemble the limitation to a specific value of  $N$  is removed and the summation can be carried out over all the occupation numbers with no hampering restriction regarding their total sum.

Thus the grand partition function can be written in a form analogous to the  $Z$  of Eqs. (19-8) and (21-10),

$$\mathcal{Q} = \sum_{\nu} g_{\nu} \exp \left[ (1/kT) \sum_j n_j (\mu - \epsilon_j) \right] \quad (24-2)$$

by virtue of Eqs. (24-1). The multiplicities  $g_{\nu}$ , for each system state  $\nu$  (i.e., for each different set of occupation numbers  $n_j$ ) are chosen according to the degree of distinguishability of the particles involved.

Indeed, this way of writing  $\mathcal{Q}$  is appropriate also when the "particles" are identical subsystems, such as the molecules of a gas. In such cases the "particle states"  $j$  are the molecular quantum states, specified by their translational, rotational, vibrational, and electronic quantum numbers, the allowed energies  $\epsilon_j$  are the sums  $\epsilon_{kmn}^{\text{tr}} + \epsilon_{\lambda\nu}^{\text{rot}}$  +  $\epsilon_n^{\text{vib}} + \epsilon^{\text{el}}$  of Chapter 21, and the  $n_j$ 's are the number of molecules that have the same totality of quantum numbers  $j \equiv k, m, n, \lambda, \nu, n$ , etc. However, we shall postpone discussion of this generalization until Chapter 27.

### Maxwell-Boltzmann Particles

At present we wish to utilize the grand canonical ensemble to investigate systems of "elementary" particles, such as electrons or protons or photons or the like, sufficiently separated so that their mutual interactions are negligible. Each particle in such a system will have the same mass  $m$  and will be subject to the same conservative forces, so that the Schrödinger equation for each will be the same. Therefore, the total set of allowed quantum numbers, represented by the index  $j$ , will be the same for each particle (although at any instant

different particles may have different values of  $j$ ). The allowed energy corresponding to the set of quantum numbers represented by  $j$  is  $\epsilon_j$  and the number of particles in this state is  $n_j$ . The grand partition function can then be written as in Eq. (24-2).

The values of the multiplicities  $g_\nu$  must now be determined for each kind of fundamental particle. This to some extent is determined by the nature of the forces acting on the particles. For example, if the particle has a spin and a corresponding magnetic moment, if no magnetic field is present the several allowed orientations of spin will have the same energy, and the  $g$ 's will reflect this fact. Let us avoid this complication at first, and assume that each particle state  $j$  has an energy  $\epsilon_j$  which differs from that of any other particle state.

When the elementary particles are distinguishable, the discussion leading up to Eq. (21-9) indicates that  $g_\nu$ , the number of different ways  $N$  particles can be assigned to the various particle states,  $n_j$  of them being in the  $j$ -th particle state, is  $[N! / n_1! n_2! \dots]$ . Since  $0! = 1$ , we can consider all the  $n$ 's as being represented in the denominator, even those for states unoccupied; an infinite product of 1's is still 1. The grand partition function will thus have the same kind of terms as in Eq. (21-10), but the restriction on the summation is now removed; all values of the  $n_j$ 's are allowed. Therefore, for *distinguishable particles*,

$$\mathcal{Z}_{\text{dist}} = \sum_{n_1, n_2, \dots} \frac{(n_1 + n_2 + \dots)!}{n_1! n_2! \dots} \exp \left[ \frac{1}{kT} \sum_j n_j (\mu - \epsilon_j) \right] \quad (24-3)$$

In contrast to Eq. (20-10), this sum is not separable into a simple product of one-particle factors; because of the lack of limitation on  $N$  it reduces to the sum  $\sum_N \left[ \sum_j e^{(\mu - \epsilon_j)/kT} \right]^N$ , each term of which is a product of one-particle partition functions.

In the classical limit, when there are many particle states in the range of energy equal to  $kT$ , the chance of two particles being in the same particle state is vanishingly small, and the preponderating terms in series (24-3) are those for which no  $n_j$  is larger than 1. In this case we may correct for the indistinguishability of particle by the "shot-gun" procedure, used in Chapter 21, of dividing every term by the total number of ways in which  $N$  particles can be arranged in  $N$  different states. The resulting partition function

$$\mathcal{Z}_{\text{MB}} = \sum_{n_1, n_2, \dots} \frac{1}{n_1! n_2! \dots} \exp \left[ (1/kT) \sum_j (\mu - \epsilon_j) \right] =$$

$$\begin{aligned}
 &= \sum_{n_1} \frac{1}{n_1!} e^{n_1(\mu - \epsilon_1)/kT} \sum_{n_2} \frac{1}{n_2!} e^{n_2(\mu - \epsilon_2)/kT} \dots \\
 &= \exp \left[ \sum_j e^{(\mu - \epsilon_j)/kT} \right] = \gamma_1 \gamma_2 \dots \\
 \gamma_j &= \exp \left[ e^{(\mu - \epsilon_j)/kT} \right] \tag{24-4}
 \end{aligned}$$

can be separated, being a product of factors  $\gamma_j$ , one for each particle state  $j$ , not one for each particle. This is the partition function we used to obtain Eqs. (23-6) and (23-7). As was demonstrated there and earlier, this way of counting states results in the thermodynamics of a perfect gas. It results also in the Maxwell-Boltzmann distribution for the mean number of particles occupying a given particle state  $j$ .

This last statement can quickly be shown by obtaining the grand potential  $\Omega$  from  $\mathcal{Q}$  and then, by partial differentiation by  $\mu$ , obtaining the mean number of particles in the system,

$$\begin{aligned}
 \Omega_{MB} &= -kT \ln (\mathcal{Q}_{MB}) = -kT \sum_j e^{(\mu - \epsilon_j)/kT} = -PV \\
 \bar{N} &= e^{\mu/kT} \sum_j e^{-\epsilon_j/kT} = \sum_j \bar{n}_j; \quad \bar{n}_j = e^{(\mu - \epsilon_j)/kT} \tag{24-5}
 \end{aligned}$$

where  $\bar{N}$  is equal to  $PV/kT$ , thus fixing the value of  $\mu = -kT \ln \left[ (kT/PV) \times \sum_j e^{-\epsilon_j/kT} \right]$ . In fact  $\mu$  acts like a magnitude parameter in the grand canonical ensemble; its value is adjusted to make  $-(\partial \Omega / \partial \mu)_{TV}$  equal to the specified value of  $\bar{N}$ . The quantity  $\bar{n}_j$ , the mean value of the occupation number for the  $j$ -th particle state for this ensemble, takes the place of the particle probabilities (for we can no longer ask what state a given particle is in). We see that the mean number of particles in state  $j$ , with energy  $\epsilon_j$ , is

$$\bar{n}_j = \left( PV/kT \sum_i e^{-\epsilon_i/kT} \right) e^{-\epsilon_j/kT} \tag{24-6}$$

which is proportional to the Maxwell-Boltzmann factor  $e^{-\epsilon_j/kT}$ .

Therefore particles that correspond to this partition function may be called *Maxwell-Boltzmann particles* (MB particles for short). No actual system of particles corresponds exactly to this distribution for all temperatures and densities. But all systems of particles approach this behavior in the limit of high-enough temperatures, whenever the classical phase-space approximation is valid.

Before the advent of the quantum theory the volume of phase space occupied by a single microstate was not known; in fact it seemed reasonable to assume that every element of phase space, no matter how small, represented a separate microstate. If this were the case, the chance that two particles would occupy the same state was of the second order in the volume element and could be neglected. Thus for classical statistical mechanics the procedure of dividing by  $N!$  was valid. Now we know that the magnitude of phase-space volume occupied by a microstate is finite, not infinitesimal; it is apparent that there can be situations in which the system points are packed closely enough in phase space so that two or more particles are within the volume that represents a single microstate; in these cases the MB statistics is not an accurate representation.

### Bosons and Fermions

Actual particles are of two types. Both types are indistinguishable and thus, according to Eq. (20-11), have multiplicity factors  $g_\nu = 1$ , rather than  $(N! / n_1! n_2! \dots)$ . A state of the system is specified by specifying the values of the occupation numbers  $n_j$ . Each such state is a single one; it has no meaning to try to distinguish which particle is in which state; all we can specify are the numbers  $n_j$  in each state.

In addition to their indistinguishability, different particles obey different rules regarding the maximum value of  $n_j$ . One set of particles can pack as many into a given particle state as the distribution will allow;  $n_j$  can take on all values from 0 to  $\infty$ . Such particles are called *bosons*; they are said to obey the *Bose-Einstein statistics* (BE for short). Photons and helium atoms are examples of bosons. For these particles the  $g_\nu$  of Eq. (24-2) are all unity and the grand partition function is

$$\begin{aligned}\mathcal{Z}_{\text{BE}} &= \sum_{n_1, n_2, \dots} \exp \left[ (1/kT) \sum_j n_j (\mu - \epsilon_j) \right] \\ &= \sum_{n_1} e^{n_1(\mu - \epsilon_1)/kT} \sum_{n_2} e^{n_2(\mu - \epsilon_2)/kT} \dots = \mathcal{Z}_1 \mathcal{Z}_2 \\ \mathcal{Z}_j &= \left[ 1 - e^{(\mu - \epsilon_j)/kT} \right]^{-1} \end{aligned}\quad (24-7)$$

where we have used Eq. (20-2) to consolidate the factor sums  $\mathcal{Z}_j$ .

Here again the grand partition function separates into factors, one for each particle state, rather than one for each particle. We note that the series for the  $j$ -th factor does not converge unless  $\mu$  is less than the corresponding energy  $\epsilon_j$ .

Here again we can calculate the grand potential and the mean number of particles in the system of bosons,

$$\Omega_{BE} = kT \sum_j \ln \left[ 1 - e^{(\mu - \epsilon_j)/kT} \right] = -PV$$

$$\bar{N} = \sum_j \bar{n}_j; \quad \bar{n}_j = \left[ e^{(\epsilon_j - \mu)/kT} - 1 \right]^{-1} \quad (24-8)$$

where  $\bar{n}_j$  is the mean number of particles in the  $j$ -th particle state. In this case there is no simple equation fixing the value of  $\mu$  in terms of  $\bar{N}$  (or of  $PV$ ) and  $T$ , nor is the relationship between  $\bar{N}$  and  $\Omega = -PV$  as simple as it was with Eq. (24-5). Nevertheless, knowing the allowed energy values  $\epsilon_j$  and the temperature  $T$ , we can adjust  $\mu$  so the sum of  $1/[e^{(\epsilon_j - \mu)/kT} - 1]$  over all  $j$  is equal to  $\bar{N}$ . This value of  $\mu$  is then used to compute the other thermodynamic quantities.

Note the difference between the occupation number  $n_j$  for the boson and that of Eq. (24-5) for the MB particle. For higher states, where  $\epsilon_j - \mu \gg kT$ , the two values do not differ much, but for the lower states, at lower temperatures, where  $\epsilon_j - \mu$  is equal to or smaller than  $kT$ , the  $n_j$  for the boson is appreciably greater than that for the MB particle (shall we call it a maxwellon?). Bosons tend to "condense" into their lower states, at low temperatures, more than do maxwellons.

The other kind of particle encountered in nature has the idiosyncrasy of refusing to occupy a state that is already occupied by another particle. In other words the occupation numbers  $n_j$  for such particles can be 0 or 1, but not greater than 1. Particles exhibiting such unsocial conduct are said to obey the Pauli exclusion principle. They are called *fermions* and are said to obey *Fermi-Dirac statistics* (FD for short). Electrons, protons, and other elementary particles with spin 1/2 are fermions. For these particles  $g_\nu = 1$ , but the sum over each  $n_j$  omits all terms with  $n_j > 1$ . Therefore,

$$\mathcal{Q}_{FD} = \sum_{n_1, n_2, \dots} \exp \left[ (1/kT) \sum_j n_j (\mu - \epsilon_j) \right] = \gamma_1 \cdot \gamma_2 \cdots$$

$$\gamma_j = \left[ 1 + e^{(\mu - \epsilon_j)/kT} \right] \quad (24-9)$$

Again the individual factors  $\gamma_j$  are for each quantum state, rather than for each particle. The mean values of the occupation numbers can be obtained from  $\mathcal{Q}$  as before,

$$\Omega_{FD} = -kT \sum_j \ln \left[ 1 + e^{(\mu - \epsilon_j)/kT} \right] = -PV$$

$$N = \sum_{j=1}^{\infty} \bar{n}_j; \quad \bar{n}_j = \left[ e^{(\epsilon_j - \mu)/kT} + 1 \right]^{-1} \quad (24-10)$$

where again the relation between  $N$  and  $PV$  (i.e., the equation of state) is not so simple as it is for Maxwellians, and again there is no simple relationship that determines  $\mu$  in terms of  $N$ ; the equation for  $\bar{N}$  must be inverted to find  $\mu$  as a function of  $\bar{N}$ .

Comparing the mean number  $\bar{n}_j$  of particles in state  $j$  for fermions with the  $\bar{n}_j$  for MB particles [Eq. (24-5)], we see that for the higher states, where  $\epsilon_j - \mu \gg kT$ , the two values are roughly equal, but for the lower states the  $\bar{n}_j$  for fermions is appreciably smaller (for a given value of  $\mu$ ) than the  $\bar{n}_j$  for Maxwellians. Fermions tend to stay away from the lower states more than do Maxwellians, and thus much more than do bosons. In fact, fermions cannot enter a state already occupied by another fermion; according to the Pauli principle  $\bar{n}_j$  cannot be larger than 1.

### Comparison among the Three Statistics

The differences between the BE, MB, and FD statistics can be most simply displayed by comparing the multiplicities  $g_j$  of the mean occupation numbers  $\bar{n}_j$ . In each case the multiplicities are products of factors  $g_j(n_j)$ , one for each particle state  $j$ . The three sets of values are

$$g_j(n_j) = \begin{cases} 1 (n_j = 0 \text{ or } 1) & = 1 \quad (n_j = 2, 3, \dots) \quad \text{BE statistics} \\ 1 (n_j = 0 \text{ or } 1) & = \frac{1}{n_j!} \quad (n_j > 1) \quad \text{MB statistics} \\ 1 (n_j = 0 \text{ or } 1) & = 0 \quad (n_j = 2, 3, \dots) \quad \text{FD statistics} \end{cases} \quad (24-11)$$

The  $g$ 's are identical for  $n_j = 0$  or 1; they differ for the higher values of the occupation numbers. Bosons have  $g_j = 1$  for all values of  $n_j$ ; they don't care how many others are in the same state. Fermions have  $g_j = 0$  for  $n_j > 1$ ; they are completely unsocial. The approximate statistics we call MB has values intermediate between 0 and 1 for  $n_j > 1$ ; these particles are moderately unsocial; the  $g_j$  tend toward zero as  $n_j$  increases.

In terms of the energy  $\epsilon_j$  of the  $j$ -th particle state and the value of the normalizing parameter  $\mu$ , the mean number of particles in state  $j$  is

$$\bar{n}_j = \begin{cases} 1 / \left[ e^{(\epsilon_j - \mu) / kT} - 1 \right] & \text{BE statistics} \\ 1 / e^{(\epsilon_j - \mu) / kT} & \text{MB statistics} \\ 1 / \left[ e^{(\epsilon_j - \mu) / kT} + 1 \right] & \text{FD statistics} \end{cases} \quad (24-12)$$

For FD statistics,  $\bar{n}_j$  can never be larger than 1; for MB statistics  $\bar{n}_j$  can be larger than 1 for those states with  $\mu$  larger than  $\epsilon_j$ ; for BE statistics  $\mu$  cannot be larger than  $\epsilon_1$  [see discussion of Eq. (24-7)] but  $\bar{n}_j$  can be much larger than 1 if  $(\epsilon_j - \mu)/kT$  is small.

In each case the value of  $\mu$  is determined by requiring that the sum of the  $\bar{n}_j$ 's, over all values of  $j$ , be equal to the mean number  $\bar{N}$  of particles in the system. If  $kT$  is large compared to the energy spacings  $\epsilon_{j+1} - \epsilon_j$ , then  $\bar{n}_{j+1}$  will not differ much from  $\bar{n}_j$  and the sum for  $\bar{N}$  will consist of a large number of  $\bar{n}_j$ 's, of slowly diminishing magnitude. Therefore much of the sum for  $\bar{N}$  will be "carried" by the  $\bar{n}_j$ 's for the higher states ( $j > 1$ ). If, at the same time,  $\bar{N}$  is small, then all the  $\bar{n}_j$ 's must be small; even  $\bar{n}_1$  must be less than 1. For this to be so,  $(\epsilon_1 - \mu)$  must be larger than  $kT$ , so that the terms  $e^{(\epsilon_j - \mu)/kT}$  must all be considerably larger than 1. In this case the values of the  $\bar{n}_j$ 's, for the three statistics, are nearly equal, and we might as well use the intermediate MB values, since these provide us with a simpler set of equations for  $\mu$ ,  $S$ ,  $P$ ,  $C_V$ , etc. [Eqs. (24-5)].

In other words, in the limit of high temperature and low density, both bosons and fermions behave like classical Maxwell-Boltzmann particles. For this reason, the fact that classical statistical mechanics is only an approximation did not become glaringly apparent until systems of relatively high density were studied at low temperatures (except in the case of photons, which are a special case).

When  $kT$  is the same size as  $\epsilon_2 - \epsilon_1$  or smaller, the three statistics display markedly different characteristics. For *bosons*  $\mu$  becomes very nearly equal to  $\epsilon_1$  ( $\mu = \epsilon_1 - \delta$ , where  $\delta \ll kT$ ) so that

$$\bar{n}_1 = \left[ e^{(\epsilon_1 - \mu)/kT} - 1 \right]^{-1} \simeq kT/\delta$$

and

$$\bar{n}_j \simeq \left[ e^{(\epsilon_j - \epsilon_1)/kT} - 1 \right]^{-1} \quad \text{for } j > 1$$

which is considerably smaller than  $\bar{n}_1$  if  $\epsilon_2 - \epsilon_1 > kT$ . Therefore at low temperatures and high densities, most of the bosons are in the lowest state ( $j = 1$ ) and

$$\bar{n}_1 \simeq N - \sum_{j=2}^{\infty} \left[ e^{(\epsilon_j - \epsilon_1)/kT} - 1 \right]^{-1} \rightarrow \bar{N}, \quad kT \rightarrow 0 \quad (24-13)$$

which serves to determine  $\delta$ , and therefore  $\mu = \epsilon_1 - \delta$ . At very low temperatures bosons "condense" into the ground state. The "condensation" is not necessarily one in space, as with the condensation of a

vapor into a liquid. The ground state may be distributed all over position space but may be "condensed" in momentum space. This will be illustrated later.

For *fermions* such a condensation is impossible; no more than one fermion can occupy a given state. As  $T \rightarrow 0$ ,  $\mu$  must approach  $\epsilon_N$ , so that  $\bar{n}_j = [e^{(\epsilon_j - \mu)/kT} + 1]^{-1}$  is practically equal to 1 for  $j < N$  (since  $e^{(\epsilon_j - \mu)/kT} \simeq e^{(\epsilon_j - \epsilon_N)/kT}$  is then very much smaller than 1) and is much smaller than 1 for  $j > N$  (since  $e^{(\epsilon_j - \mu)/kT}$  is then very large compared to 1 for  $\epsilon_j > \epsilon_N$ ). Thus at low temperatures the lowest  $N$  particle states are completely filled with fermions (one per state) and the states above this "Fermi level"  $\epsilon_N$  are devoid of particles.

The behavior of *MB particles* differs from that of either bosons or fermions at low temperatures and high densities. The lower states are populated by more than one particle, in contrast to the fermions, but they don't condense exclusively and suddenly in just the ground state, as do bosons. The comparison between the number of particles per unit energy range, for a gas of bosons, one of fermions and one of maxwellons, is shown in Fig. 24-1. We see that fermions pack the lower  $N$  levels solidly but uniformly, that bosons tend to concentrate in the very lowest state, and that maxwellons are intermediate in distribution. Because of the marked difference in behavior from that of a classical perfect gas, a gas of bosons or fermions at low temperatures and high densities is said to be *degenerate*.

### Distribution Functions and Fluctuations

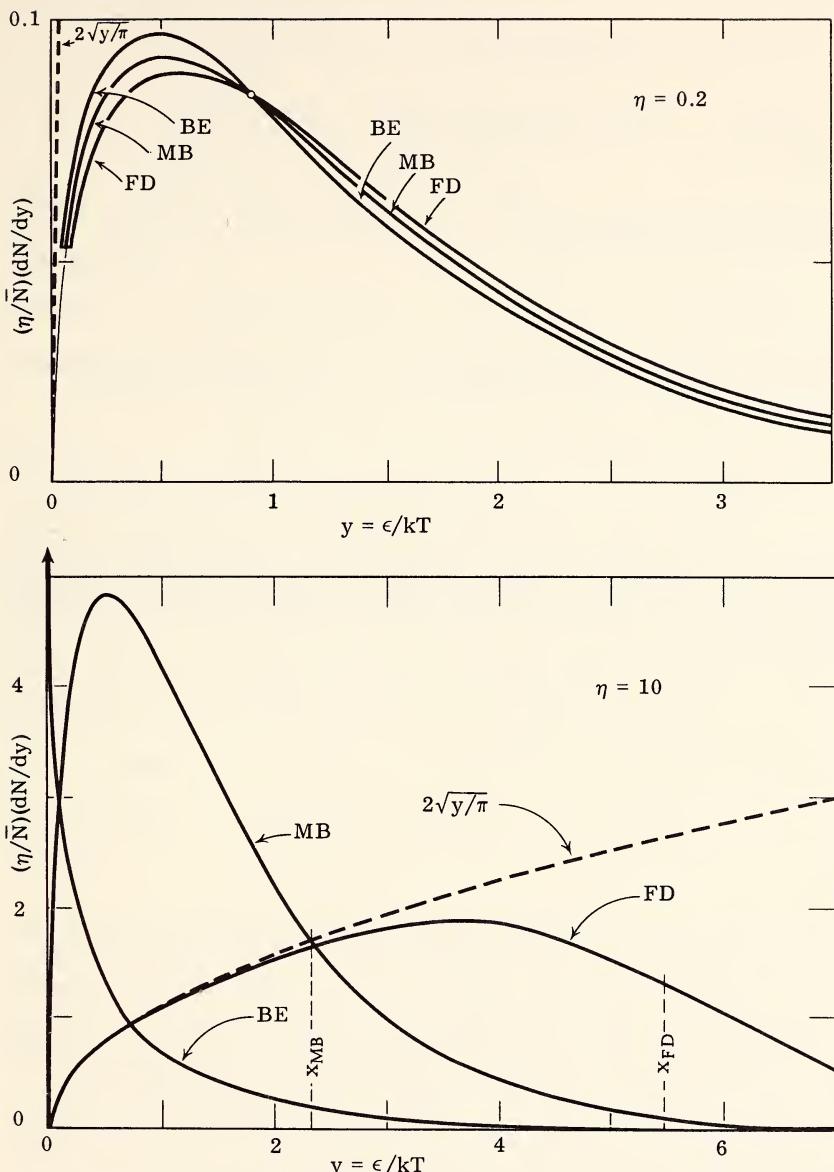
With indistinguishable particles there is no sense in asking the probability that a specific particle is in state  $j$ ; all we can ask for is the probability  $f_j(n_j)$  that  $n_j$  particles are in state  $j$ . These probabilities can be obtained from the distribution function of the ensemble, given in Eq. (23-4). For

$$f_{N\nu} = (g_\nu/2) \exp \left[ \sum_{j=1}^{\infty} n_j (\mu - \epsilon_j)/kT \right] = f_1(n_1) \cdot f_2(n_2) \cdots$$

$$f_j(n_j) = (g_j/2_j) e^{n_j (\mu - \epsilon_j)/kT} \quad (24-14)$$

where the factor  $2_j$  of the partition function, for the  $j$ -th particle state, is given by Eq. (24-4), (24-7), or (24-9), depending upon whether the particles in the system are MB, BE, or FD particles.

To be specific, the probability that  $n$  particles are in state  $j$  (we can leave the subscript off  $n$  without producing confusion here) for the three statistics, is



**Fig. 24-1.** Density of particles per energy range for a gas, according to the three statistics, for nondegenerate and degenerate conditions. Dashed curve  $2(y/\pi)^{1/2}$  corresponds to a density of one particle per particle state. Area under each curve equals  $\eta$ . See also page 233.

$$\text{For bosons: } f_j(n) = e^{n(\mu - \epsilon_j)/kT} - e^{(n+1)(\mu - \epsilon_j)/kT}$$

$$\text{For MB particles: } f_j(n) = \frac{1}{n!} \exp \left[ n \frac{\mu - \epsilon_j}{kT} - e^{(\mu - \epsilon_j)/kT} \right]$$

$$\text{For fermions: } f_j(n) = \begin{cases} e^{n(\mu - \epsilon_j)/kT} / \left[ 1 + e^{(\mu - \epsilon_j)/kT} \right] & n = 0, 1 \\ 0 & n > 1 \end{cases}$$

(24-15)

Reference to Eqs. (24-12) shows that the mean value of  $n$  is given by the usual formula,

$$\bar{n}_j = \sum_{n=0}^{\infty} n f_j(n) \quad (24-16)$$

With a bit of algebraic juggling, we can then express the probability  $f_j(n)$  in terms of  $n$  and of its mean value  $\bar{n}_j$  (we can call it  $\bar{n}$  without confusion here):

$$f_j(n) = \begin{cases} (\bar{n})^n / (\bar{n} + 1)^{n+1} & \text{for bosons} \\ [(\bar{n})^n / n!] e^{-\bar{n}} & \text{for MB particles} \\ 1 - \bar{n} \text{ if } n = 0, = \bar{n} \text{ if } n = 1, = 0 \text{ if } n > 1 & \text{for fermions} \end{cases} \quad (24-17)$$

The distribution function  $f_j(n)$  for bosons is a *geometric* distribution. The ratio  $f_j(n)/f_j(n-1)$  is a constant,  $\bar{n}/(\bar{n} + 1)$ ; the chance of adding one more particle to state  $j$  is the same, no matter how many bosons are already in the state. The MB distribution is the familiar Poisson distribution of Eqs. (11-15) and (23-10), with ratio  $f_j(n)/f_j(n-1) = \bar{n}/n$ , which decreases as  $n$  increases. The presence of Maxwellians in a given state discourages the addition of others, to some extent. On the other hand,  $f_j(n)$  for FD statistics is zero for  $n > 1$ ; if a fermion occupies a given state, no other particle can join it (the Pauli principle).

Using these expressions for  $f_j(n)$  we can calculate the variance  $(\Delta n_j)^2$  of the occupation number  $n_j$  for state  $j$ , for each kind of statistics:

$$(\Delta n_j)^2 = \sum_{n=0}^{\infty} (n - \bar{n}_j)^2 f_j(n) = \sum_n n^2 f_j(n) - (\bar{n}_j)^2 =$$

$$= \begin{cases} \bar{n}_j(\bar{n}_j + 1) & \text{for bosons} \\ \bar{n}_j & \text{for MB particles} \\ \bar{n}_j(1 - \bar{n}_j) & \text{for fermions} \end{cases} \quad (24-18)$$

and, from this, obtain the fractional fluctuation  $\Delta n_j/\bar{n}_j$  of the occupation numbers,

$$\Delta n_j/\bar{n}_j = \begin{cases} \sqrt{1 + (1/\bar{n}_j)} & \text{for bosons} \\ \sqrt{1/\bar{n}_j} & \text{for MB particles} \\ \sqrt{(1/\bar{n}_j) - 1} & \text{for fermions} \end{cases} \quad (24-19)$$

The fractional fluctuation is greatest for the least-occupied states ( $\bar{n}_j \ll 1$ ). As the mean occupation number increases the fluctuation decreases, going to zero for fermions as  $\bar{n}_j \rightarrow 1$  (the degenerate state) and to zero for Maxwellians as  $\bar{n}_j \rightarrow \infty$ . But the standard deviation  $\Delta n_j$  for bosons is never less than the mean occupancy  $\bar{n}_j$ . We shall see later that the local fluctuations in intensity of thermal radiation (photons are bosons) are always large, of the order of magnitude of the intensity itself, as predicted by Eq. (24-19).

# 25

## Bose-Einstein Statistics

The previous chapter has indicated that, as the temperature is lowered or the density is increased, systems of bosons or of fermions enter a state of degeneracy, wherein their thermodynamic properties differ considerably from those of the corresponding classical system, subject to Maxwell-Boltzmann statistics. These differences are apparent even when the systems are perfect gases, where the interaction between particles is limited to the few collisions needed to bring the gas to equilibrium. Indeed, in some respects, the differences between the three statistics are more apparent when the systems are perfect gases than when they are more complex in structure. Therefore it is useful to return once again to the system we started to study in Section 2, this time to analyze in detail the differences caused by differences in statistics. In this chapter we take up the properties of a gas of bosons. Two different cases will be considered; a gas of photons (electromagnetic radiation) and a gas of material particles, such as helium atoms.

### General Properties of a Boson Gas

Using Eqs. (24-8) et seq., we compute the distribution function, mean occupation numbers, and thermodynamic functions for the gas of bosons:

$$f_j(n) = \left[ 1 - e^{(\mu - \epsilon_j)/kT} \right] e^{n(\mu - \epsilon_j)/kT} = \left[ \bar{n}_j^n / (\bar{n}_j + 1)^{n+1} \right]$$

$$\bar{n}_j = \sum_{n=0}^{\infty} n f_j(n) = \left[ e^{(\epsilon_j - \mu)/kT} - 1 \right]^{-1}$$

$$\Omega_{BE} = -PV = -kT \ln \mathcal{Q} = kT \sum_{j=1}^{\infty} \ln \left[ 1 - e^{(\mu - \epsilon_j)/kT} \right]$$

$$\bar{N} = -(\partial\Omega/\partial\mu)_{TV} = \sum_{j=1}^{\infty} \bar{n}_j = \sum_{j=1}^{\infty} \left[ e^{(\epsilon_j - \mu)/kT} - 1 \right]^{-1}$$

$$S = -(\partial\Omega/\partial T)_{\mu V} = (U - \bar{N}\mu - \Omega)/T$$

$$U = \sum_{j=1}^{\infty} \epsilon_j \bar{n}_j; \quad C_V = T(\partial S/\partial T)_{\mu V} \quad (25-1)$$

where  $\mu$  must be less than the lowest particle energy  $\epsilon_1$  in order that the series expansions converge. All these quantities are functions of the chemical potential  $\mu$ . For systems in which the mean number of particles  $\bar{N}$  is specified, the value of  $\mu$ , as a function of  $\bar{N}$ ,  $V$ , and  $T$ , is determined implicitly by the equation for  $\bar{N}$  given above. The value obtained by inverting this equation is then inserted in the other equations, to give  $S$ ,  $U$ ,  $P$ , and  $C_V$  as functions of  $\bar{N}$ ,  $V$ , and  $T$ .

In the case of the photon gas, in equilibrium at temperature  $T$  in a volume  $V$  (black-body radiation), the number of photons  $\bar{N}$  in volume  $V$  is not arbitrarily specified; it adjusts itself so that the radiation is in equilibrium with the constant-temperature walls of the container. Since, at constant  $T$  and  $V$ , the Helmholtz function  $F = \Omega + \mu\bar{N}$  comes to a minimal value at equilibrium [see the discussion following Eq. (8-10)], if  $\bar{N}$  is to be varied to reach equilibrium at constant  $T$  and  $V$ , we must have

$$(\partial F/\partial \bar{N})_{TV} = \frac{\partial}{\partial \bar{N}} (\Omega + \mu\bar{N}) = \mu \text{ equal to zero} \quad (25-2)$$

Therefore, for a photon gas at equilibrium, at constant  $T$  and  $V$ , the chemical potential of the photons must be zero [see the discussion following Eq. (7-8)].

### Classical Statistics of Black-Body Radiation

At this point the disadvantages of a "logical" presentation of the subject become evident; a historical presentation would bring out more vividly the way experimental findings forced a revision of classical statistics. It was the work of Planck, in trying to explain the frequency distribution of electromagnetic radiation which first exhibited the inadequacy of the Maxwell-Boltzmann statistics and pointed the way to the development of quantum statistics. A purely logical demonstration, that quantum statistics does conform with observation, leaves out the atmosphere of struggle which permeated the early development of quantum theory, struggle to find a theory that would fit the many new and unexpected measurements.

Experimentally, the energy density of black-body radiation having frequency between  $\omega/2\pi$  and  $(\omega + d\omega)/2\pi$  was found to fit an empirical formula

$$d\epsilon = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3 d\omega}{e^{\hbar\omega/kT} - 1} \rightarrow \begin{cases} (\omega^2 kT/\pi^2 c^3) d\omega & kT \gg \hbar\omega \\ (\hbar\omega^3/\pi^2 c^3)e^{-\hbar\omega/kT} d\omega & kT \ll \hbar\omega \end{cases}$$

where, at the time,  $\hbar$  was an empirical constant, adjusted to fit the formula to the experimental curves. Classical statistical mechanics could explain the low-frequency part of the curve ( $kT \gg \hbar\omega$ ) but could not explain the high-frequency part (Fig. 25-1).

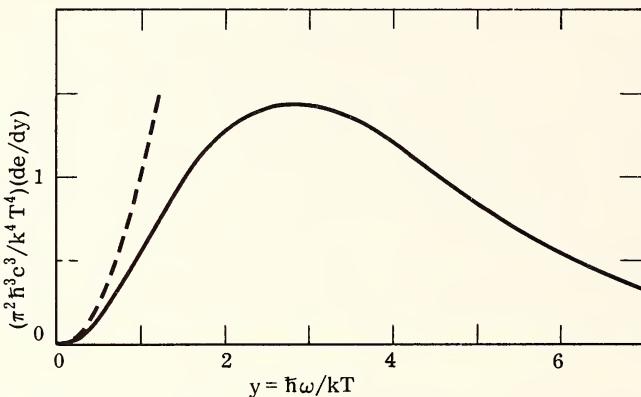


Fig. 25-1. The Planck distribution of energy density of black-body radiation per frequency range. Dashed line is Rayleigh-Jeans distribution.

Classically, each degree of freedom of the electromagnetic radiation should possess a mean energy  $kT$  [see the discussion of Eq. (15-5)], so determining the formula for  $d\epsilon$  should simply involve finding the number of degrees of freedom of the radiation between  $\omega$  and  $\omega + d\omega$ . Since the radiation is a collection of standing waves, it can proceed exactly as was done in Chapter 20, in finding the number of standing waves in a crystal with frequencies between  $\omega/2\pi$  and  $(\omega + d\omega)/2\pi$  [see Eqs. (20-10) and (20-11)]. In a rectangular enclosure of sides  $l_x, l_y, l_z$  the allowed values of  $\omega$  are

$$\omega_j = \pi c [(k/l_x)^2 + (m/l_y)^2 + (n/l_z)^2]^{1/2} \quad (25-4)$$

where  $k, m, n$  are integers and where  $c$  is the velocity of light. Each different combination of  $k, m, n$  corresponds to a different elec-

tromagnetic wave, a different degree of freedom or, in quantum language, a different quantum state  $j$  for a photon.

By methods completely analogous to those used in Chapter 20, we find that the number of different degrees of freedom having allowed values of  $\omega_j$  between  $\omega$  and  $\omega + d\omega$  are

$$dj = (V/\pi^2 c^3) \omega^2 d\omega, \quad V = l_x l_y l_z \quad (25-5)$$

which is twice the value given in Eq. (20-11) because light can have two mutually perpendicular polarizations, so there are two different standing waves for each set of values of  $k$ ,  $m$ , and  $n$ . As mentioned before, this formula is valid for nonrectangular enclosures of volume  $V$ .

Now, if each degree of freedom carries a mean energy  $kT$ , then the total energy within  $V$ , between  $\omega$  and  $\omega + d\omega$ , is  $(kT) dj$  and the energy density of radiation with frequency between  $\omega/2\pi$  and  $(\omega + d\omega)/2\pi$  is

$$de = (kT/V) dj = (\omega^2 kT / \pi^2 c^3) d\omega$$

which is called the Rayleigh-Jeans formula. We see that it fits the empirical formula (25-3) at the low-frequency end (see the dashed curve of Fig. 25-1) but not for high frequencies.

As a matter of fact it is evident that the Rayleigh-Jeans formula cannot hold over the whole range of  $\omega$  from 0 to  $\infty$ , for the integral of  $de$  would then diverge. If this were the correct formula for the energy density then, to reach equilibrium with its surroundings, a container filled with radiation would have to withdraw an infinite amount of energy from its surroundings; all the thermal energy in the universe would drain off into high-frequency electromagnetic radiation. This outcome was dramatized by calling it the *ultraviolet catastrophe*. There is no sign of such a fate, so the Rayleigh-Jeans formula cannot be correct for high frequencies. In fact the empirical curve has the energy density  $de$  dropping down exponentially, according to the factor  $e^{-\hbar\omega/kT}$ , when  $\hbar\omega \gg kT$ , so that the integral of the empirical expression does not diverge.

Parenthetically, a similar catastrophe cannot arise with waves in a crystal because a crystal is not a continuous medium; there can only be as many different standing waves in a crystal as there are atoms in the crystal; integration over  $\omega$  only goes to  $\omega_m$  [see Eq. (26-13)] not to  $\infty$ . In contrast, the electromagnetic field is continuous, not atomic, so there is no lower limit to wavelength, no upper limit to the frequency of its standing waves.

A satisfactory exposition (to physicists, at any rate) would be to proceed from empirical formula (25-3) to the theoretical model that

fits it, showing that the experimental findings lead inexorably to the conclusion that photons obey Bose-Einstein statistics. We have not the space to do this; we shall show instead that assuming photons are bosons (with  $\mu = 0$ ) leads directly to the empirical formula (25-3) and by identifying the empirical constant  $\hbar = h/2\pi$  with Planck's constant, joins the theory of black-body radiation to all the rest of quantum theory.

### Statistical Mechanics of a Photon Gas

As we have already pointed out in Eq. (25-2), photons are a rather special kind of boson; their chemical potential is zero when they are in thermal equilibrium in volume  $V$  at temperature  $T$ . Formulas (25-1) thus simplify. For example, the mean number of photons in state  $j$  is  $\bar{n}_j = (e^{\epsilon_j/kT} - 1)^{-1}$ . But state  $j$  has been defined as the state that has frequency  $\omega_j/2\pi$ , where  $\omega_j$  is given in Eq. (25-4) in terms of its quantum numbers. Since a photon of frequency  $\omega_j/2\pi$  has energy  $\epsilon_j = \hbar\omega_j$ , the mean occupation number becomes

$$\bar{n}_j = 1 / (e^{\hbar\omega_j/kT} - 1) \quad (25-6)$$

Since there are  $(V/\pi^2c^3)\omega^2 d\omega = dj$  different photon states (different standing waves) with frequencies between  $\omega/2\pi$  and  $(\omega+d\omega)/2\pi$ , the mean number of photons in this frequency range in the container is

$$dn = \frac{V}{\pi^2 c^3} \frac{\omega^2 d\omega}{e^{\hbar\omega/kT} - 1} \quad (25-7)$$

The mean energy density  $de$  of black-body radiation in this frequency range is  $dn$  times the energy  $\hbar\omega$  per photon, divided by  $V$ , which turns out to be identical with the empirical formula for  $de$  given in Eq. (25-3). Thus the assumption that photons are bosons with  $\mu = 0$  leads directly to agreement with observation.

The frequency distribution of radiation given in Eq. (25-3) is called the *Planck distribution*. The energy density per unit frequency band increases proportional to  $\omega^2$  at low frequencies; it has a maximum at  $\omega = 2.82(kT/\hbar)$  [where  $x = 2.82$  is the solution of the equation  $(3-x)e^x = 3$ ] and it drops exponentially to zero as  $\omega$  increases beyond this maximum. Measurements have checked all these details; in fact this was the first way by which the value of  $h$  was determined. The mean number of photons, and the mean energy density, of all frequencies can be obtained from the following formulas:

$$\int_0^\infty \frac{x^2 dx}{e^x - 1} = 2.404; \quad \int_0^\infty \frac{x^3 dx}{e^x - 1} = \frac{\pi^4}{15} = 6.494 \quad (25-8)$$

For example, the mean energy density is

$$e(T) = \int de = \frac{(kT)^4}{\pi^2 c^3 \hbar^3} \int_0^\infty \frac{x^3 dx}{e^x - 1} = a T^4, \quad a = \frac{\pi^2 k^4}{15 \hbar^3 c^3} \quad (25-9)$$

which is the same as Eq. (7-8), of course, only now we have obtained an expression for Stefan's constant  $a$  in terms of  $k$ ,  $h$ , and  $c$  (which checks with experiment).

The grand potential  $\Omega$  (which also equals  $F$ , since  $\mu = 0$ ) is

$$\begin{aligned} \Omega &= kT \int dn \ln \left( 1 - e^{-\hbar\omega/kT} \right) = \frac{V k T}{\pi^2 c^3} \int_0^\Omega \omega^2 d\omega \ln \left( 1 - e^{-\hbar\omega/kT} \right) \\ &= -\frac{k^4 T^4 V}{3 \pi^2 c^3 \hbar^3} \int_0^\infty \frac{x^3 dx}{e^x - 1} = -\frac{1}{3} a V T^4 = -\frac{1}{3} V e(T) = F \end{aligned} \quad (25-10)$$

where we have integrated by parts. The other thermodynamic quantities are obtained by differentiation,

$$S = -\left(\frac{\partial F}{\partial T}\right)_V = \frac{4}{3} a V T^3; \quad P = -\left(\frac{\partial F}{\partial V}\right)_T = \frac{1}{3} a T^4 = \frac{1}{3} e(T) \quad (25-11)$$

which also check with Eqs. (7-8). The mean number of photons of any frequency in volume  $V$  is

$$\bar{N} = \int dn = \frac{V k^3 T^3}{\pi^2 c^3 \hbar^3} \int_0^\infty \frac{x^2 dx}{e^x - 1} = \frac{2.404}{\pi^2} \left(\frac{kT}{c\hbar}\right)^3 V = \frac{0.625 k^3}{c\hbar} V T^3 \quad (25-12)$$

### Statistical Properties of a Photon Gas

We saw earlier that the assumption of classical equipartition of energy for each degree of freedom of black-body radiation leads to the nonsensical conclusion that each container of radiation has an infinite heat capacity. The assumption that photons are bosons, with  $\mu = 0$ , leads to the Planck formula, rather than the Rayleigh-Jeans formula, and leads to the conclusion that the mean energy carried per degree of freedom of the thermal radiation is

$$\hbar\omega \bar{n}_j = \left( \hbar\omega_j / \left( e^{\hbar\omega_j/kT} - 1 \right) \right) \rightarrow \begin{cases} \frac{kT}{\hbar\omega_j e^{-\hbar\omega_j/kT}} & \hbar\omega_j \ll kT \\ \hbar\omega_j & \hbar\omega_j \gg kT \end{cases} \quad (25-13)$$

which equals  $kT$ , the classical value, only for low-frequency radiation. At high frequencies each photon carries so much energy that, even in thermal equilibrium, very few can be excited (just as with other quantized oscillators that have an energy level spacing large compared to  $kT$ ), and the mean energy possessed by these degrees of freedom falls off proportionally to  $e^{-\hbar\omega/kT}$ . Thus, as we have seen, the mean energy  $e(T)$  is not infinite as classical statistics had predicted.

As was pointed out at the end of Chapter 24, the fluctuations in a boson gas are larger than those in a classical gas. For a photon gas the standard deviation  $\Delta n_j$  of the number of photons in a particular state  $j$ , above and below the mean value  $\bar{n}_j$ , is  $\sqrt{\bar{n}_j(\bar{n}_j + 1)}$  and consequently the fractional fluctuation is

$$\Delta n_j/\bar{n}_j = \sqrt{(\bar{n}_j + 1)/\bar{n}_j} = e^{\hbar\omega/2kT} \quad (25-14)$$

This is also equal to the fractional fluctuation of energy density  $\Delta e_j/e_j$  or of intensity  $\Delta I_j/I_j$  of the standing wave having frequency  $\omega_j/2\pi$ . This quantity is always greater than unity, indicating that the standard deviation of the intensity of a standing wave is equal to or greater than its mean intensity.

Such large fluctuations may be unusual for material gases; they are to be expected for standing waves. If the  $j$ -th wave is fairly steadily excited (i.e., if  $\bar{n}_j > 1$ , i.e., if  $e^{\hbar\omega_j/kT} < 2$ ) then it will be oscillating more or less sinusoidally and its intensity will vary more or less regularly between zero and twice its mean value, which corresponds to a standard deviation roughly equal to its mean value. If, on the other hand, the standing wave is excited only occasionally, the sinusoidal oscillation will occur only occasionally and the amplitude will be zero in between times. In this case the standard deviation will be larger than the mean. Thus Eq. (25-14) is not as anomalous as it might appear at first.

The variance  $(\Delta n)^2$  of the number of photons in all the standing waves in volume  $V$  having frequency between  $\omega/2\pi$  and  $(\omega + d\omega)/2\pi$  is the sum of the variances of the component states,

$$(\Delta n)^2 = (\Delta n_j)^2 \left( \frac{V\omega^2 d\omega}{\pi^2 c^3} \right) = \frac{V\omega^2 d\omega}{\pi^2 c^3} \frac{e^{\hbar\omega/kT}}{\left( e^{\hbar\omega/kT} - 1 \right)^2}$$

and the fractional fluctuation of energy in this frequency range is

$$\Delta dn/dn = \sqrt{\frac{V\omega^2 d\omega}{\pi^2 c^3}} e^{\hbar\omega/2kT} \quad (25-15)$$

The wider the frequency band  $d\omega$  and the greater the volume consid-

ered, the smaller is the fractional fluctuation; including a number of standing waves in the sample "smooths out" the fluctuations.

### Statistical Mechanics of a Boson Gas

When the bosons comprising the gas are material particles, rather than photons,  $\mu$  is not zero but is determined by the mean particle density. The particle energy  $\epsilon$  is not  $\hbar\omega$  but is the kinetic energy  $p^2/2m$  of the particle if  $m$  is its mass. We have already shown [see Eqs. (19-2) and (21-7)] that, for elementary particles in a box of "normal" size, the translational levels are spaced closely enough so that we can integrate over phase space instead of summing over particle states. The number of particle states in an element  $dx dy dz dp_x \times dp_y dp_z = dV_q dV_p$  of phase space is  $g(dV_q dV_p/h^3)$  where  $g$  is the multiplicity factor caused by the particle spin. If the spin is  $s$  and no magnetic field is present,  $g = (2s + 1)$  different spin orientations have the same energy  $\epsilon$ . Therefore the sum for  $\bar{N}$  of Eq. (25-1) becomes

$$\begin{aligned}\bar{N} &= (g/h^3) \int \dots \int \left[ e^{(\epsilon - \mu)/kT} - 1 \right]^{-1} dV_q dV_p \\ &= (gV/h^3) \int_0^{2\pi} d\beta \int_0^\pi \sin \alpha \, d\alpha \int_0^\infty \left[ e^{(\epsilon - \mu)/kT} - 1 \right]^{-1} p^2 \, dp\end{aligned}\quad (25-16)$$

where angles  $\alpha$  and  $\beta$  are the spherical angles in momentum space of Eq. (12-1).

We can integrate over  $\alpha$  and  $\beta$  and, since  $\epsilon = (p^2/2m)$  or  $p = \sqrt{2m\epsilon}$ , we can change to  $\epsilon$  for the other integration variable, so

$$\begin{aligned}\bar{N} &= 2\pi gV \left( \frac{2m}{h^2} \right)^{3/2} \int_0^\infty \frac{\sqrt{\epsilon} \, d\epsilon}{e^{(\epsilon - \mu)/kT} - 1} \\ &= gV \left( \frac{2\pi m kT}{h^2} \right)^{3/2} f_{1/2}(-\mu/kT)\end{aligned}\quad (25-17)$$

where

$$f_m(x) = \frac{1}{m!} \int_0^\infty \frac{z^m \, dz}{e^{z+x} - 1} = \sum_{n=1}^{\infty} \left( e^{-nx} / n^{m+1} \right) \rightarrow e^{-x}, \quad x \rightarrow \infty$$

The series for  $f_m$  converges if  $x$  is positive. However we recollect that with Bose-Einstein statistics  $\mu$  must be less than the lowest energy level, which is zero for gas particles. Therefore  $\mu$  is negative and  $x = -(\mu/kT)$  is positive, and the series does converge.

It should be pointed out that the change from summation to integration has one defect; it leaves out the ground state  $\epsilon = 0$ . This term, in the sum of particle states, is the largest term of all; in the integral approximation it is completely left out because the density function  $\sqrt{\epsilon}$  goes to zero there. Ordinarily this does not matter, for there are so many terms in the sum for  $\bar{N}$  for  $\epsilon$  small compared to  $kT$  (which are included in the integral) that the omission of this one term makes a negligible difference in the result. At low temperatures, however, bosons "condense" into this lowest state [see Eq. (24-13)] and its population becomes much greater than that for any other state. We shall find that above a limiting temperature  $T_0$  the ground state is no more densely populated than many of its neighbors and that it can then be neglected without damage. Below  $T_0$ , however, the lowest state begins to collect more than a normal share of particles, and we have to add an extra term to the integral for  $\bar{N}$ , corresponding to the number of particles that have "condensed" into the zero-energy state.

We should have mentioned this complication when we were discussing a photon gas, of course, for the integrals of (25-9) to (25-11) also have left out the zero-energy state. But a photon of zero energy has zero frequency, so this lowest energy state represents a static electromagnetic field. We do not usually consider a static field as an assemblage of photons and, furthermore, the exact number of photons present is not usually of interest; the measurable quantities are the energy density and intensity. For more-material bosons, however, the mean number of particles can be measured directly, so we must account for the excess of particles in the zero-energy state when the gas is degenerate.

### Thermal Properties of a Boson Gas

The value of  $-\mu$  is determined implicitly by the equation

$$\eta \equiv \bar{N}\ell_t^3/gV = f_{1/2}(x); \quad x = -\mu/kT; \quad \ell_t = h/\sqrt{2\pi mkT} \quad (25-18)$$

which can be inverted to obtain  $-\mu$  as a function of  $T$  and  $\eta$ . When the parameter  $\eta$  is small (low density and/or high temperature),  $f_{1/2}$  has its limiting form  $e^{-x} = e^{\mu/kT}$  and

$$\mu \rightarrow -kT \ln(gV/\bar{N}\ell_t^3) = kT \ln \eta, \quad \eta \rightarrow 0$$

which is the value for a classical, perfect gas of Eqs. (23-7). The computed values of  $x$  for larger values of  $\eta$  are given in Table 25-1.

The internal energy  $U$  and the grand potential  $\Omega$  of the gas can also be expressed as integrals,

TABLE 25-1  
Functions for a Boson Gas

$\eta$	$x$	$T/T_0$	$PV/\bar{N}kT$	$S/\bar{N}k$	$2C_V/3\bar{N}k$	$N_C/\bar{N}$
0	$\infty$	$\infty$	1.000	$\infty$	1.00	0
0.1	2.342	8.803	0.977	4.784	1.01	0
1	0.358	1.897	0.818	2.403	1.09	0
2	0.033	1.195	0.637	1.625	1.19	0
2.5	0.001	1.030	0.536	1.341	1.26	0
2.612	0	1.000	0.513	1.282	1.28	0
3	0	0.912	0.447	1.116	1.12	0.129
10	0	0.409	0.134	0.335	0.33	0.739
30	0	0.196	0.045	0.112	0.11	0.913
$\infty$	0	0	0	0	0	1.000

$$U = 2\pi gV(2m/h^2)^{3/2} \int_0^\infty \frac{\epsilon^{3/2} d\epsilon}{e^{(\epsilon - \mu)/kT} - 1} = \frac{3}{2} (\bar{N}kT/\eta) f_{3/2}(-\mu/kT)$$

$$\begin{aligned} -\Omega = +PV &= -2\pi kTgV(2m/h^2)^{3/2} \int_0^\Omega \sqrt{\epsilon} d\epsilon \ln \left[ 1 - e^{(\mu - \epsilon)/kT} \right] \\ &= \frac{2}{3} U \end{aligned}$$

$$S = -(\partial\Omega/\partial T)_{V\mu} = \frac{5}{3}(U/T) + (-\mu\bar{N}/T) = \bar{N}k \left[ x + \frac{5}{2} \frac{f_{3/2}(x)}{f_{1/2}(x)} \right] \quad (25-19)$$

where we have integrated the expression for  $-\Omega$  by parts to obtain  $(2/3)U$ , and where we have used the equation  $(d/dx)f_m(x) = -f_{m-1}(x)$  to obtain the expression for the entropy  $S$ . Compare these formulas with the ones of Eqs. (21-14) for a gas of MB particles. The term  $x$  here corresponds to the logarithmic terms in the formula for  $S$ , for example, and the two expressions for  $S$  are identical when  $f_{3/2} = f_{1/2}$ , i.e., when  $\eta \rightarrow 0$ .

The heat capacity  $C_V$  could be computed by taking the partial of  $U$  with respect to  $T$ , holding  $\bar{N}$  and  $V$  constant. But the independent variables here are  $\mu$ ,  $T$ , and  $V$  and, rather than bothering to change variables, we use the formula  $C_V = T(\partial S/\partial T)_{\mu V}$ . The values can be computed by numerical differentiation or by series expansion. The important quantities are tabulated in Table 25-1 for a range of values of the parameter  $\eta$ . We see that when  $\eta$  is small,  $PV$  is practically equal to  $\bar{N}kT$  (the perfect gas law) and  $C_V$  is practically equal to  $(3/2)\bar{N}k$ .

Equations (25-17) and (25-19) show that

$$(S/\bar{N}k) = \frac{5}{2} [f_{3/2}(-\mu/kT)/f_{1/2}(-\mu/kT)] - (\mu/kT) \quad \text{and}$$

$$(\bar{N}/VT^{3/2}) = g(2\pi mk/h^2)^{3/2} f_{1/2}(-\mu/kT)$$

are functions of  $-\mu/kT$  alone. In an adiabatic expansion both  $\bar{N}$  and  $S$  remain constant; therefore in an adiabatic expansion  $-\mu/kT$  and  $VT^{3/2}$  stay constant for a boson gas. We can also show that  $PV^{5/3}$  is constant during an adiabatic expansion. These results are identical with Eqs. (4-12) and (6-5) for a classical perfect gas. Evidently a boson gas behaves like a perfect gas of point particles in regard to adiabatic compression, although its equation of state is not that of a perfect gas (Table 25-1 shows that  $PV/\bar{N}kT$  diminishes as  $\eta$  increases).

When  $\eta$  is less than unity, a first-order approximation is

$$\begin{aligned} -\Omega = PV &= \frac{2}{3}U \simeq \bar{N}kT [1 - (\eta/2^{5/2})] \\ &\simeq \bar{N}kT \left[ 1 - \frac{\bar{N}}{2gV} \left( \frac{h^2}{4\pi mkT} \right)^{3/2} \right] \end{aligned} \quad (25-20)$$

from which we can obtain  $C_V$  by differentiation of  $U$  (since the independent variables are now  $\bar{N}$ ,  $T$ , and  $V$ ). The boson gas exhibits a smaller pressure and a larger specific heat than a classical perfect gas, at least for moderate temperatures and densities.

### The Degenerate Boson Gas

As the density of particles is increased and/or the temperature is decreased,  $\eta$  increases,  $x = -\mu/kT$  decreases and the thermal properties of the gas depart farther and farther from those of a classical perfect gas, until at  $\eta = 2.612$ ,  $\mu$  becomes zero. If  $\eta$  becomes larger than this, Eq. (25-17) no longer can be satisfied. For the maximum value of  $f_{1/2}(x)$  is 2.612, for  $\mu = 0$ , and  $\mu$  cannot become positive. The only way the additional particles can be accommodated is to put them into the hitherto-neglected zero-energy state mentioned several pages back.

If  $\bar{N}$  is held constant and  $T$  is reduced, the condensation starts when  $\eta = 2.612$ , and thus when  $T$  reaches the value

$$T_0 = (h^2/2\pi mk)(\bar{N}/2.612gV)^{2/3} = 3.31(\hbar^2/mk)(\bar{N}/gV)^{2/3} \quad (25-21)$$

Any further reduction of  $T$  will force some of the particles to condense into the zero-energy state. In fact the number  $N_X$  of particles that can stay in the upper states are those which satisfy Eq. (25-17) with  $\mu = 0$ .

$$N_X = 2.612gV(2\pi mkT/h^2)^{3/2} = \bar{N}(T/T_0)^{3/2} \quad (25-22)$$

and the rest,

$$N_C = \bar{N}[1 - (T/T_0)^{3/2}]$$

are condensed in the ground state, exerting no pressure and carrying no energy. Therefore, the thermodynamic functions for the gas in this partly condensed state are

$$PV = -\Omega = \frac{2}{3}U = 0.513\bar{N}kT(T/T_0)^{3/2} = 0.086 \frac{m^{3/2}gV}{h^3} (kT)^{5/2}$$

or

$$P = 0.086(m^{3/2}g/h^3)(kT)^{5/2}$$

$$S = 5U/3T = 1.28\bar{N}k(T/T_0)^{3/2} = \frac{2}{3}C_V \quad (25-23)$$

The pressure is independent of volume, because this is all the pressure the uncondensed particles can withstand. Further reduction of volume simply condenses more particles into the ground state, where they contribute nothing to the pressure. The heat capacity of the gas as a function of  $T$  has a discontinuity in slope at  $T_0$ , as shown in Fig. 25.2. At high temperatures the gas is similar to an MB gas of point particles, with  $C_V = (3/2)\bar{N}k$ . As  $T$  is diminished  $C_V$  rises until, at  $T = T_0$ , it has its largest value,  $C_V = 1.92\bar{N}k$ . For still smaller values of  $T$ ,  $C_V$  decreases rapidly, to become zero at  $T = 0$ .

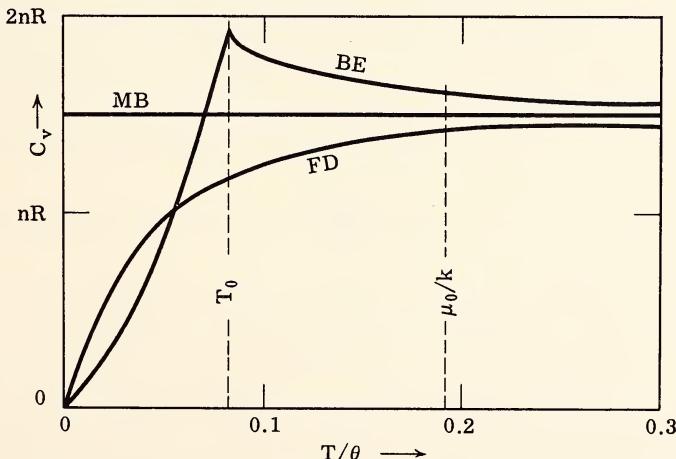


Fig. 25-2. Heat capacity of a gas (according to the three statistics) versus temperature in units of  $\theta = (h^2/mk)(\bar{N}/gV)^{2/3}$ .

The "condensed" particles are not condensed in position space, as in change of phase; they are "condensed" in momentum space, at  $p = 0$ , a set of  $\bar{N}[1 - (T/T_0)^{3/2}]$  stationary particles, distributed at random throughout the volume  $V$ . Liquid helium II acts like a mixture of a condensed phase (superfluid) plus some ordinary liquid, the fraction of superfluid increasing as  $T$  is decreased. Since HeII is a liquid the theoretical model to account for its idiosyncrasies is much more complicated than the gas formulas we have developed here. Although the theory is not yet complete, the assumption that  $\text{He}^4$  atoms are bosons does explain many of the peculiar properties of HeII.

# 26

## Fermi-Dirac Statistics

Fermi-Dirac statistics is appropriate for electrons and other elementary particles that are subject to the Pauli exclusion principle. The occupation numbers  $n_j$  can only be zero or unity and the mean number of particles in state  $j$  is the  $\bar{n}_j$  of Eq. (24-10). In this chapter we shall work out the thermal properties of a gas of fermions, to compare with those of a gas of bosons and with those of a perfect gas of MB particles, particularly in the region of degeneracy. There are no FD analogues to photons, with  $\mu = 0$ .

### General Properties of a Fermion Gas

For a gas of fermions, at temperature  $T$  in a volume  $V$ , the particle energy is  $\epsilon = p^2/2m$  as before, and the number of allowed translational states in the element of phase space  $dV_q dV_p$  is  $g(dV_q dV_p/h^3)$  as before ( $g$  is the spin multiplicity  $2s + 1$ ). Integrating over  $dV_q$  and over all directions of the momentum vector, we find the number of states with kinetic energy between  $\epsilon$  and  $\epsilon + d\epsilon$  is  $2\pi gV(2m/h^2)^{3/2}\sqrt{\epsilon} d\epsilon$ , as before. Multiplying this by  $\bar{n}_j$  [Eq. (24-10)] gives us the mean number of fermions with kinetic energy between  $\epsilon$  and  $\epsilon + d\epsilon$ ,

$$dN = 2\pi gV(2m/h^2)^{3/2} \frac{\sqrt{\epsilon} d\epsilon}{e^{(\epsilon - \mu)/kT} + 1} \quad (26-1)$$

which is to be compared with the integrand of Eq. (25-17) for the boson gas and with  $dN = 2\pi gV(2m/h^2)^{3/2} e^{(\mu - \epsilon)/kT} \sqrt{\epsilon} d\epsilon$  for a perfect gas of MB particles.

Figure 24-1 compares plots of  $dN/d\epsilon$  for these three statistics for two different degrees of degeneracy. As  $\eta = (\bar{N}/gV)(h^2/2\pi mkT)^{3/2}$  varies, the MB distribution changes scale, but not shape. For small values of  $\eta$ , the values of  $\mu = -xkT$  for the three cases are all negative and do not differ much in value, nor do the three curves differ much in shape. In this region the MB approximation is satisfactory.

For large values of  $\eta$  the curves differ considerably, and the values of the chemical potential  $\mu$  differ greatly for the three cases. For bosons, as we saw in Chapter 25,  $\mu$  is zero and a part of the gas has "condensed" into the ground state, making no contribution to the energy or pressure of the gas, and being represented on the plot by the vertical arrow at  $y = 0$ . For fermions,  $\mu$  is positive, and the states with  $\epsilon$  less than  $\mu$  are practically completely filled, whereas those with  $\epsilon$  greater than  $\mu$  are nearly empty. Because of the Pauli principle, no more than one particle can occupy a state; at low temperatures and/or high densities the lowest states are filled, up to the energy  $\epsilon = \mu$ , and the states for  $\epsilon > \mu$  are nearly empty, as shown by the curve (which is the parabolic curve  $2V\sqrt{\epsilon}/\pi$  for  $y$  less than  $-x$  and which drops to zero shortly thereafter).

The dotted parabola  $2V\sqrt{\epsilon}/\pi$  corresponds to the level density  $dN/d\epsilon = 2\pi gV(2m/h^2)^{3/2} \sqrt{\epsilon}$ , corresponding to one particle per state. We see that the curve for MB particles rises above this for  $\eta$  large, corresponding to the fact that some of the lower levels have more than one particle apiece. The BE curve climbs still higher at low energies. The FD curve, however, has to keep below the parabola everywhere.

The conduction electrons in a metal are the most accessible example of a fermion gas. In spite of the fact that these electrons are moving through a dense lattice of ions, they behave in many respects as though the lattice were not present. Their energy distribution is more complicated than the simple curves of Fig. 24-1 and, because of the electric forces between them and the lattice ions, the pressure they exert on the container is much less than that exerted by a true gas; nevertheless their heat capacity, entropy, and mean energy are remarkably close to the Fermi-gas values. Measurements on conduction electrons constitute most of the verifications of the theoretical model to be described in this chapter.

### The Degenerate Fermion Gas

As  $T$  approaches zero the FD distribution takes on its fully degenerate form, with all states up to the  $\bar{N}$ -th completely filled and all states beyond the  $\bar{N}$ -th completely empty. In other words, the limiting value of  $\mu$  (call it  $\mu_0$ ) is large and positive, and

$$dN = \begin{cases} 2\pi gV(2m/h^2)^{3/2} \sqrt{\epsilon} \, d\epsilon & 0 \leq \epsilon \leq \mu_0 \\ 0 & \epsilon > \mu_0 \end{cases} \quad (26-2)$$

where  $\mu_0$  has the value that allows the integral of  $dN$  to equal  $\bar{N}$ ,

$$\bar{N} = 2\pi gV(2m/h^2)^{3/2} \int_0^{\mu_0} \sqrt{\epsilon} \, d\epsilon \quad \text{or} \quad \mu_0 = \beta(\bar{N}/V)^{2/3} \quad (26-3)$$

where  $\beta = (h^2/2m)(3/4\pi g)^{2/3} = 5.84 \times 10^{-38}$  joule-meter<sup>2</sup> for electrons ( $g = 2$ ).

Even at absolute zero most of the fermions are in motion, some of them moving quite rapidly. For an electron gas of density  $\rho = \bar{N}m/V$  kg per m<sup>3</sup>, the kinetic energy of the fastest,  $\mu_0$ , is roughly equal to  $40\rho^{2/3}$  electron volts;  $\mu_0/k$  is approximately equal to  $4.5 \times 10^5 \rho^{2/3}$  °K. In other words the top of the occupied levels (the Fermi level) corresponds to the mean energy [=  $(3/2)kT$ ] of a MB particle in a gas at the temperature  $3 \times 10^5 \rho^{2/3}$  °K. For the conduction electrons in metals, where  $\rho_{el} \approx 3 \times 10^{-2}$  kg/m<sup>3</sup>, this corresponds to about 30,000°K; for free electrons in a white-dwarf star, where  $\rho_{el} > 1000$ , it corresponds to more than  $3 \times 10^7$  degrees. Until the actual temperature of a Fermi gas is larger than this value, it remains degenerate. The parameter  $\eta = 0.752(\mu_0/kT)^{3/2}$  is a good index of the onset of degeneracy (when  $\eta > 1$  there is degeneracy).

The internal energy of the completely degenerate gas (which, like the boson gas, is equal to  $-(3/2)\Omega$  at all temperatures), is

$$U_0 = \int_0^{\mu_0} \epsilon dN = \frac{3}{5} \beta \bar{N} (\bar{N}/V)^{2/3} = \frac{3}{5} \bar{N} \mu_0 = -\frac{3}{2} \Omega \quad (26-4)$$

$$P_0 = -\Omega/V = \frac{2}{5} \beta (\bar{N}/V)^{5/3}; \quad S_0 = 0$$

Even at absolute zero a fermion gas exerts pressure. If electrons were neutral particles, their pressure would be about  $2.7 \times 10^7 \rho^{5/3}$  atmospheres at absolute zero. Because of the strong electrical attractions to the ions of the crystal lattice, this pressure is largely counterbalanced by the forces holding the crystal together.

When  $T$  is small compared to  $\mu_0/k$  (i.e., when  $\eta$  is larger than unity) but is not zero, a first-order correction to the formulas for complete degeneracy can be worked out. The results are

$$U \approx U_0 + \frac{1}{4} \pi^2 (\bar{N}/\mu_0) (kT)^2 = \bar{N} \mu_0 \left[ \frac{3}{5} + \frac{1}{4} \pi^2 (kT/\mu_0)^2 \right]$$

$$F \approx U_0 - \frac{1}{4} \pi^2 (\bar{N}/\mu_0) (kT)^2 = \frac{3}{5} \beta \bar{N}^{5/3} V^{-2/3} - \frac{\pi^2}{4\beta} (kT)^2 \bar{N}^{1/3} V^{2/3}$$

$$S \approx \frac{1}{2} \pi^2 \bar{N} k (kT/\mu_0) \approx C_p \approx C_v \approx \frac{1}{2} \pi^2 (k^2 T/\beta) \bar{N}^{1/3} V^{2/3}$$

$$P \approx \frac{2}{5} \beta (\bar{N}/V)^{5/3} + (\pi^2/6\beta) (kT)^2 (\bar{N}/V)^{1/3} \quad (26-5)$$

These formulas verify that, as long as  $T$  is small compared with

$\mu_0/k$ , the fermion gas is degenerate, with thermal properties very different from those of a classical, perfect gas. The internal energy is nearly constant, instead of being proportional to  $T$ ; the pressure is inversely proportional to the  $5/3$  power of the volume and its dependence on  $T$  is small.

The heat capacity of the degenerate gas is proportional to  $T$  at low temperatures, being considerably smaller than the classical value  $3Nk$  when  $T$  is less than  $\mu_0/k$ . Thus the  $C_V$  of the conduction electrons is small compared to the lattice  $C_V$  for metals at room temperatures. However, the heat capacity of the lattice of ions is proportional to  $T^3$  for low temperatures [see Eq. (20-16)] so that if  $T$  is made small enough, the linear term, for the conduction electron gas, will predominate over the cubic term for the lattice. It is found experimentally that, below about  $3^{\circ}\text{K}$  the heat capacity of metals is linear in  $T$  instead of cubic, as are nonconductors. This experimental fact was one of the first verifications of the theoretical prediction (made by Sommerfeld) that the conduction electrons in a metal behave like a degenerate Fermi gas.

### Behavior at Intermediate Temperatures

When  $T$  is considerably larger than  $\mu_0/k$ , the FD gas is no longer degenerate; it has roughly the same properties as the MB gas. For example, its equation of state at these high temperatures is

$$PV \simeq NkT \left[ 1 + \left( \eta / 2^{5/2} \right) \right] \simeq NkT \left[ 1 + \frac{1}{3\sqrt{2\pi}} (\mu_0/kT)^{3/2} \right] \quad (26-6)$$

which differs from the corresponding result for the boson gas of Eq. (25-20) only by the difference in sign inside the brackets. The pressure is somewhat greater than that for a perfect gas; the effect of the Pauli exclusion principle is similar to that of a repulsive force between the particles.

For intermediate temperatures the thermodynamic properties must be computed numerically. Referring to Eq. (26-1), we define a parameter  $\eta$ , as with Eq. (25-18),

$$\begin{aligned} \eta(x) &= \frac{\bar{N}}{gV} \left( \frac{\hbar^2}{2\pi mkT} \right)^{3/2} = \frac{4}{3\sqrt{\pi}} \left( \frac{\mu_0}{kT} \right)^{3/2} = \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{\sqrt{u} du}{e^{u+x} + 1} \\ &\equiv F_{1/2}(x) \rightarrow e^{-x}, \quad x \rightarrow \infty \end{aligned} \quad (26-7)$$

where  $x = -(\mu/kT)$  can be considered to be a function of  $\eta$ . The other thermodynamic quantities, being functions of  $x$ , are therefore functions of  $\eta$ ,

$$-\Omega = PV = \frac{2}{3}U = NkT(F_{3/2}/\eta)$$

where

$$F_{3/2} \equiv \frac{4}{3\sqrt{\pi}} \int_0^{\infty} \frac{u^{3/2} du}{e^{u+x} + 1} \rightarrow e^{-x}, \quad x \rightarrow \infty$$

$$S = Nk \left[ \frac{5}{3} (F_{3/2}/F_{1/2}) + x \right]; \quad F = -NkT [x + (F_{3/2}/F_{1/2})]$$

$$x = -(\mu/kT) \quad (26-8)$$

Values of some of these quantities for a few values of the density parameter  $\eta$  are given in Table 26-1. The onset of degeneracy corresponds roughly to  $\eta = 1$ . For  $\eta < 1$ ,  $PV$  is practically equal to  $NkT$  and  $C_V$  nearly equal to  $(3/2)Nk$ ; the gas is a perfect gas. When  $\eta > 1$ ,  $\mu$  is positive,  $PV$  is much larger than  $NkT$ ,  $S$  goes to zero, and  $C_V$  is much smaller than  $(3/2)Nk$ ; the gas is degenerate. The curve for  $C_V$  is shown in Fig. 25-2, in comparison with those for a perfect gas and a boson gas. These numbers should be compared with those of Table 25-1, for the boson gas.

TABLE 26-1  
Functions for a Fermion Gas

$\eta$	$x$	$kT/\mu_0$	$PV/NkT$	$S/Nk$	$2C_V/3Nk$
0	$\infty$	$\infty$	1.000	$\infty$	1.000
0.01	4.60	17.81	1.001	7.1	0.997
0.1	2.26	3.841	1.017	4.8	0.989
1	-0.35	0.827	1.174	2.6	0.919
10	-5.46	0.178	2.521	0.85	0.529
100	-26.0	0.038	10.48	0.18	0.145
316	-56.0	0.008	22.48	0.09	0.084
$\infty$	$-\infty$	0	$\infty$	0	0

# 27

## Quantum Statistics for Complex Systems

This chapter is a mixed bag. In the first part we discuss the way one can work out the statistical properties of systems that are more complex than the simple gases studied in the preceding chapters. Here we show why helium atoms can behave like elementary Bose-Einstein particles, for example, and why there has to be a symmetry factor  $\sigma$  in Eq. (22-4).

### Wave Functions and Statistics

We cannot go much further in discussing quantum statistics without talking about wave functions. As any text on quantum mechanics will state, a wave function is a solution of a Schrödinger equation; its square is a probability density. For a single particle of mass  $m$  in a potential field  $\phi(\mathbf{r})$  the equation is  $H\Psi = \epsilon\Psi$ , where  $H$  is the differential operator,

$$H \equiv -\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + \phi(\mathbf{r}) \quad (27-1)$$

which is applied to the wave function  $\Psi(\mathbf{r})$ . The values of the energy factor  $\epsilon$ , for which the equation can be solved to obtain a continuous, single-valued, and finite  $\Psi$ , are the allowed energies  $\epsilon_j$  for the particle. The square of the corresponding solutions  $\Psi_j(\mathbf{r})$  (the square of its magnitude, if  $\Psi$  is complex) is equal to the probability density that the particle is at the point  $\mathbf{r}$ . Therefore  $\Psi$  must be *normalized*,

$$\iiint |\Psi_j(\mathbf{r})|^2 dx dy dz = 1 \quad (27-2)$$

The mathematical theory of such equations easily proves that wave functions for different states  $i$  and  $j$  are *orthogonal*,

$$\iiint \bar{\Psi}_i(\mathbf{r}) \Psi_j(\mathbf{r}) dx dy dz = 0 \quad \text{unless } i = j \quad (27-3)$$

The wave function  $\Psi_j(\mathbf{r})$  embodies what we can know about the particle in state  $j$ . According to quantum theory, we cannot know the particle's exact position; so we cannot expect to obtain a solution of its classical motion by finding  $x$ ,  $y$ , and  $z$  as functions of time. All we can expect to obtain is the probability that the particle is at  $\mathbf{r}$  at time  $t$ , which is  $|\Psi|^2$ . The relation between classical and quantum mechanics is the relation between the operator  $H$  of Eq. (27-1) and the Hamiltonian function  $H(q,p)$  of Eqs. (13-9) and (16-4). For a single particle (the  $k$ -th one, say)

$$H_k(q,p) = \frac{1}{2m} \left( p_{kx}^2 + p_{ky}^2 + p_{kz}^2 \right) + \phi(\mathbf{r}_k)$$

We see that the quantum-mechanical operator is formed from the classical Hamiltonian by substituting  $(\hbar/i)(\partial/\partial q)$  for each  $p$ . For this reason we call the  $H$  of Eq. (27-1) a *Hamiltonian operator*.

The generalization to a system of  $N$  similar particles is obvious. If there is no interaction between the particles, the Hamiltonian for the system is the sum of the single-particle Hamiltonians,

$$H(p,q) = \sum_{k=1}^N H_k(q,p)$$

and the Schrödinger equation for the system is

$$H\Psi = E\Psi; \quad H = \sum_{k=1}^N H_k; \quad H_k = -\frac{\hbar^2}{2m} \nabla_k^2 + \phi(\mathbf{r}_k) \quad (27-4)$$

where

$$\nabla_k^2 = \frac{\partial^2}{\partial x_k^2} + \frac{\partial^2}{\partial y_k^2} + \frac{\partial^2}{\partial z_k^2}$$

The values of  $E$  for which there is a continuous, single-valued, and finite solution  $\Psi_\nu(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$  of Eq. (27-4) are the allowed values  $E_\nu$  of the energy of the system. They are, of course, the sums of the single-particle energies  $\epsilon_j$ , one for each particle. We have used these facts in previous chapters [see Eqs. (19-1) and (23-1), for example].

A possible solution of Eq. (27-4) is a simple product of single-particle wave functions,

$$\begin{aligned} \Psi_\nu(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) &= \Psi_{j_1}(\mathbf{r}_1) \cdot \Psi_{j_2}(\mathbf{r}_2) \cdots \Psi_{j_N}(\mathbf{r}_N) \\ E_\nu &= \sum_{k=1}^N \epsilon_{j_k} \end{aligned} \quad (27-5)$$

where  $j_k$  stands for the set of quantum numbers of the  $k$ -th particle. This would be an appropriate solution for distinguishable particles, for it has specified the state of each particle; state  $j_1$  for particle 1, state  $j_2$  for particle 2, and so on. The square of  $\Psi_\nu$  is a product of single-particle probability densities  $|\Psi_{j_k}(r_k)|^2$  that particle  $k$ , which is in the  $j_k$ -th state, is at  $r_k$ . We should note that for particles with spin, each  $\Psi$  has a separate factor which is a function of the spin coordinate, a different function for each different spin state. Thus coordinate  $r_k$  represents not only the position of the particle but also its spin coordinate, and the quantum numbers represented by  $j_k$  include the spin quantum number for the particle.

### Symmetric Wave Functions

This product wave function, however, will not do for indistinguishable particles. What is needed for them is a probability density that will have the same value if particle 1 is placed at point  $r$  (including spin) as it has if particle  $k$  is placed there. To be more precise, we wish a probability density  $|\Psi_\nu(r_1, r_2, \dots, r_N)|^2$  which is *unchanged* in value when we interchange the positions (and spins) of particle 1 and particle  $k$ , or any other pair of particles. The simple product wave function of Eq. (27-5) does not provide this; if  $j_1$  differs from  $j_k$ , then interchanging  $r_1$  and  $r_k$  produces a different function.

However, other solutions of Eq. (27-4), having the same value  $E_\nu$  of the energy of the system as does solution (27-5), can be obtained by interchanging quantum numbers and particles. For example,

$$\Psi_\nu(r_N, r_{N-1}, \dots, r_1) = \Psi_{j_N}(r_1) \cdot \Psi_{j_{N-1}}(r_2) \dots \Psi_{j_1}(r_N)$$

is another solution with energy  $E_\nu$ . There are  $N!$  possible permutations of  $N$  different quantum numbers among  $N$  different particle wave functions. If several different particles have the same quantum numbers, if  $n_j$  particles are in state  $j$ , for example, then there are  $(N! / n_1! n_2! \dots)$  [compare with Eq. (21-9)] different wave functions  $\Psi_\nu$  which can be obtained from (27-5) by permuting quantum numbers and particles.

Therefore a possible solution of Eq. (27-4), for the allowed energy  $E_\nu$ , would be a *sum* of all the different product functions that can be formed by permuting states  $j$  among particles  $k$ . Use of Eqs. (27-2) and (27-3) can show that for such a sum to be normalized, it must be multiplied by  $\sqrt{n_1! n_2! \dots / N!}$ . However, such details need not distract us here; what is important is that this sum is a solution of Eq. (27-4) for the system state  $\nu$  with energy  $E_\nu$ , which is unchanged in value when any pair of particle coordinates is interchanged (the change rearranges the order of functions in the sum but does not introduce new

terms). Therefore, its square is unchanged by such interchange and the wave function is an appropriate one for indistinguishable particles. For such a wave function it is no longer possible to talk about *the* state of a particle; all particles participate in all states; all we can say is that  $n_j$  particles are in state  $j$  at any time. Such a wave function is said to be *symmetric* to interchange of particle coordinates.

A few examples of symmetric wave functions for two particles are  $\Psi_1(\mathbf{r}_1)\Psi_1(\mathbf{r}_2)$  and  $(1/\sqrt{2})[\Psi_1(\mathbf{r}_1)\Psi_2(\mathbf{r}_2) + \Psi_1(\mathbf{r}_2)\Psi_2(\mathbf{r}_1)]$ ; a few for three particles are  $\Psi_1(\mathbf{r}_1)\Psi_1(\mathbf{r}_2)\Psi_1(\mathbf{r}_3)$  or  $(1/\sqrt{3})[\Psi_1(\mathbf{r}_1)\Psi_1(\mathbf{r}_2)\Psi_2(\mathbf{r}_3) + \Psi_1(\mathbf{r}_1)\Psi_1(\mathbf{r}_3)\Psi_2(\mathbf{r}_2) + \Psi_1(\mathbf{r}_3)\Psi_1(\mathbf{r}_2)\Psi_2(\mathbf{r}_1)]$ ; and so on.

### Antisymmetric Wave Functions

However, since our basic requirement is that of symmetry for the *square* of  $\Psi$ , we have an alternative choice, that of picking a wave function *antisymmetric* with respect to interchange of particle coordinates, which changes its sign but not its magnitude when the coordinates of any pair are interchanged. The square of such a  $\Psi$  also is unchanged by the interchange. Such an antisymmetric solution can be formed out of the product solutions of Eq. (27-5), but *only* if all particle  $\Psi$ 's are for *different states*. If every  $j_k$  differs from every other  $j$ , then an antisymmetric solution of Eq. (27-4), with energy  $E_\nu$ , is the determinant

$$\Psi_\nu = \frac{1}{\sqrt{N!}} \begin{vmatrix} \Psi_{j_1}(\mathbf{r}_1) & \Psi_{j_1}(\mathbf{r}_2) & \dots & \dots & \dots & \Psi_{j_1}(\mathbf{r}_N) \\ \Psi_{j_2}(\mathbf{r}_1) & \Psi_{j_2}(\mathbf{r}_2) & \dots & \dots & \dots & \Psi_{j_2}(\mathbf{r}_N) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \Psi_{j_N}(\mathbf{r}_1) & \Psi_{j_N}(\mathbf{r}_2) & \dots & \dots & \dots & \Psi_{j_N}(\mathbf{r}_N) \end{vmatrix} \quad (27-6)$$

The properties of determinants are such that an interchange of any two columns (interchange of particle coordinates) or of any two rows (interchange of quantum numbers) changes the sign of  $\Psi_\nu$ . The proof that  $1/\sqrt{N!}$  must be used to normalize this function is immaterial here. What is important is that another whole set of wave functions, satisfying the requirements of particle indistinguishability, is the set of functions that are antisymmetric to interchange of particle coordinates. For this set, no state can be used more than once (a determinant with two rows identical is zero).

By now it should be apparent that the two types of wave functions correspond to the two types of quantum statistics. Wave functions for a system of *bosons* are *symmetric* to interchange of particle coordi-

nates; any number of particles can occupy a given particle state. Wave functions for *fermions* are *antisymmetric* to interchange of particle coordinates; because of the antisymmetry, no two particles can occupy the same particle state (which is the Pauli exclusion principle). Both sets of wave functions satisfy the indistinguishability requirement—that the square of  $\Psi$  be symmetric to interchange of particle coordinates. A simple application of quantum theory will prove that no system of forces which are the same for all particles will change a symmetric wave function into an antisymmetric one, or vice versa. Once a boson, always a boson, and likewise for fermions. It is an interesting indication of the way that all parts of quantum theory “hang together” that the fact that the quantity of physical importance is the *square* of the wave function should not only allow, but indeed *demand*, two different kinds of statistics, one for symmetric wave functions, the other for antisymmetric.

We have introduced the subject of symmetry of wave functions by talking about a system with no interactions between particles. But this restriction is not binding. Suppose we have a system of  $N$  particles, having a classical Hamiltonian  $H(q, p)$  which includes potential energies of interaction  $\phi(r_{kl})$ , which is symmetric to interchange of particles (as it must be if the particles are identical in behavior). Corresponding to this  $H$  is a Hamiltonian operator  $H$ , which includes the interaction terms  $\phi(r_{kl})$  and in which each  $p_k$  is changed to  $(\hbar/i) \times (\partial/\partial q_k)$ ; this operator is also symmetric to interchange of particle coordinates. There are then two separate sets of solutions of the general Schrödinger equation

$$H\Psi(r_1, r_2, \dots, r_N) = E\Psi$$

One set is symmetric to interchange of coordinates,

$$\Psi(r_1, \dots, r_j, \dots, r_\ell, \dots, r_N) = \Psi(r_1, \dots, r_\ell, \dots, r_j, \dots, r_N)$$

and the other is antisymmetric,

$$\Psi(r_1, \dots, r_k, \dots, r_\ell, \dots, r_N) = -\Psi(r_1, \dots, r_\ell, \dots, r_k, \dots, r_N)$$

When there are particle interactions, the two sets usually have different allowed energies  $E_\nu$ ; in any case a function of one set cannot change into one of the other set.

The symmetric set represents a system of bosons and the antisymmetric set represents a system of fermions. It is by following through the requirements of symmetry of the wave functions that we can work out the idiosyncrasies of systems at low temperatures and high densities. By this means we can work out the thermal properties of systems with strong interactions, mentioned early in Chapter 24. At high tem-

peratures and/or low densities, Maxwell-Boltzmann statistics is adequate and we need not concern ourselves as to whether the wave function is symmetric or antisymmetric.

### Wave Functions and Equilibrium States

If a system is made up of  $N$  particles of one kind and  $M$  of another, the combined wave function of the whole is a product of an appropriately symmetrized function of the  $N$  particles, times another for the  $M$  particles, with its proper symmetry. Thus, for a gas of  $N$  hydrogen atoms, the complete wave function would be an antisymmetric function of all the electronic coordinates (including spin) times another antisymmetric for all the protons (including their spins); for both electrons and protons are fermions.

This completely antisymmetrized wave function is appropriate for the state of final equilibrium, when we have waited long enough so that electrons have exchanged positions with other electrons, from one atom to the other, so that any electron is as likely to be around one proton as another. However, it takes a long time for electrons to interchange protons in a rarefied gas. Measurements are usually made in a shorter time, and correspond to an intermediate "metastable" equilibrium, in which atoms as a whole change places, but electrons do not interchange protons. A wave function for such a metastable equilibrium is one made up of separate hydrogen atomic wave functions, symmetrized with regard to interchanging atoms as individual subsystems.

Thus it is usually more appropriate to express the wave function for a system of molecules in terms of each molecule as a separate subsystem, taking into account the effects of exchanging molecule with molecule as individual units, and arranging it as though the individual particles in one molecule never interchange with those in another. The total wave function is thus put together out of products of molecular wave functions, according to the appropriate symmetry for interchange of whole molecules, each molecular wave function organized in regard to the appropriate symmetry for exchange of like particles *within* the molecule.

Actually the differentiation between "metastable" and "long-term" equilibrium in a gas is academic in all but a few cases. By the time the temperature has risen beyond the boiling point of most gases, so that the system *is* a gas, both the translational and rotational motions of the molecules can be treated classically, as was shown in Chapters 21 and 22, and questions of symmetry no longer play a role. Only a few cases exist where the relation between wave-function symmetry and statistics is apparent. These cases each involved puzzling discrepancies with the familiar classical statistics; their resolution constituted one of the major vindications of the new statistics.

One such case is for helium, gas and liquid. Normal helium ( $\text{He}^4$ ) is made up of a nucleus of two protons and two neutrons, surrounded by two electrons. In the bound nuclear state the spins of each heavy-particle pair are opposed, so that the net spin of the  $\text{He}^4$  nucleus is zero, as is the net spin of the electrons in the lowest electronic state. Since protons, neutrons, and electrons are each fermions, the combined wave function for, say, two helium atoms could be a product of three antisymmetric functions, one for all four electrons, another for four protons, and a third for the four neutrons. This wave function would correspond to "long-term" equilibrium, since it includes the possibility of interchange of neutrons, for example, between the two nuclei. A more realistic wave function would be formed of products of separate atomic wave functions.

For example, we could assume that electrons 1 and 2 were in atom a, electrons 3 and 4 in atom b, and similarly with the neutrons and protons. The electronic wave function for atom a would then be antisymmetric for interchange of electrons 1 and 2, that for atom b antisymmetric for interchange of 3 and 4. We would not consider interchanging 1 and 3 or 1 and 4 separately, only interchanging atom a as a whole with atom b. Since interchanging atoms interchanges two electrons, and two protons and two neutrons, the effect of the interchange would be symmetric, since  $(-1)^2 = +1$ . Therefore the system wave function should be symmetric for interchange of atoms;  $\text{He}^4$  atoms should behave like bosons. As we mentioned at the end of Chapter 25, liquid  $\text{He}^4$  does exhibit "condensation" effects of the sort predicted for BE particles at low temperatures. In contrast  $\text{He}^3$ , which has only one neutron in its nucleus instead of two, and thus should not behave like a boson, does not exhibit condensation effects.

Other molecules also should behave like bosons (those with an even number of elementary particles) but they become solids before they have a chance to exhibit the effects of degeneracy. Normal helium ( $\text{He}^4$ ) is the only substance with small enough interatomic forces to be still fluid at temperatures low enough for boson condensation to take place; and once started, the condensation prevents solidification down to (presumably) absolute zero. At pressures above 25 atm  $\text{He}^4$  does solidify, and none of the condensation effects are noticeable.

### Electrons in a Metal

Another system in which the effects of quantum statistics are noticeable is the metallic crystal. In the case of nonmetallic crystals the usual assumption is valid—that each atom acts as a unit and that there is not sufficient time (during most experiments) for the atoms to change places. Consequently the Debye theory of crystal heat capacities does not consider the consequences of indistinguishability of the constituent atoms. In the case of magnetic effects, however, where

the spins of some electrons are more strongly coupled to their counterparts on other atoms than to their neighbors in the same atom, the symmetry of the spin parts of these wave functions must be considered.

In the case of metals, the more tightly bound electrons, in the inner shells of the lattice ions, do not readily move from ion to ion. But the outer electrons can change places with their neighbors easily. Consequently the complete wave function for the metallic crystal can be written as a product of individual ionic wave functions, one for each ion in the lattice (with appropriate symmetry within each ionic factor) times an antisymmetric function for all the conduction electrons. The individual electron wave functions  $\Psi_j(\mathbf{r}_k)$  are not very similar to the standing waves of free particles [see Eq. (21-5)]; after all, they are traveling through a crystal lattice, not in force-free space. But they can exchange with their neighbors rapidly enough so that it is not possible to specify which electron is on which ion.

The combined wave function for the conduction electrons thus must be an antisymmetric combination like that of Eq. (27-6), which means that the usual type of FD degeneracy will take place over about the same range of temperatures (0 to about  $1000^\circ$ ) as if these electrons were free particles in a gas. Since the thermal properties of a degenerate Fermi gas depend more on the degeneracy than on the exact form of the wave functions, these conduction electrons behave more like a pure Fermi gas than anyone expected (until it was worked out by Sommerfeld). In their electrical properties, of course, the effects of their interaction with the lattice ions becomes important; but even here the effects of degeneracy are still controlling.

### Ortho- and Parahydrogen

As mentioned a few pages ago, a system composed of hydrogen molecules ( $\text{H}_2$ ) behaves like an MB gas as far as its translational energy goes. Each molecule behaves as though it is an indivisible subsystem and the whole wave function can be considered to be a product of single-molecule wave functions, each of them being in turn products of translational, rotational, vibrational, and electronic wave functions, corresponding to the separation of energies of Eq. (21-1). Each of these molecular factors of course must have the symmetry or antisymmetry required by the particles composing the molecule. For example, because protons are fermions, interchange of the two protons composing the molecule must change the sign of the wave function. (This is *not* the case with the HD molecule, where the two nuclei are dissimilar.)

Each proton has a position, relative to the center of mass of the molecule, and a spin. If the two spins are opposed, so that the total nuclear spin of the molecule is zero (singlet state) the spin part of the

nuclear wave function is antisymmetric. Therefore the space part of the nuclear wave function must be symmetric, in order that the product of both will change sign when the two are interchanged. On the other hand if the spins are parallel, so that the total nuclear spin is 1 (triplet state) the spin factor is symmetric and the space part of the nuclear wave function must be antisymmetric. Now the factor in the molecular wave function which achieves the interchange of position of the nuclei is the rotational factor, which is a spherical harmonic of the angles denoting the direction of the axis through the nuclei. Interchanging the positions of the nuclei corresponds to rotating the axis through  $180^\circ$ .

As mentioned in connection with Eq. (22-1) the allowed values of the square of the angular momentum of the molecule are  $\hbar^2 \ell(\ell + 1)$ , where  $\ell$  is the order of the spherical harmonic in the corresponding rotational wave function. It turns out that those spherical harmonics with even values of  $\ell$  are symmetric with respect to reversing the direction of the molecular axis; those with odd values of  $\ell$  change sign when the axis is rotated  $180^\circ$ . The upshot of all this is that those  $H_2$  molecules which are in the singlet nuclear state, with opposed spins, can only have even values of  $\ell$  and those in the triplet nuclear state can have only odd values of  $\ell$ .

The nuclear spins are well protected from outside influences, and it is an exceedingly rare collision which disturbs them enough for the spin of one molecule to affect that of another. Therefore, the molecule that is in the singlet nuclear state stays for days in the singlet state and is practically a different molecule from one in a triplet nuclear state. Only after a long time (or in the presence of special catalysts) do the spins exchange from molecule to molecule, so that the whole system finally comes to over-all, long-term equilibrium.

The two kinds of  $H_2$  are permanent enough to be given different names; the singlet type, with even values of  $\ell$ , is called *parahydrogen*. Its rotational partition function is

$$z_{\text{rot}}^p = \sum_{\ell=0,2,4, \dots} (2\ell + 1) \exp[-\theta_{\text{rot}} \ell(\ell + 1)/T] \quad (27-7)$$

instead of the sum over all values of  $\ell$  as in Eq. (22-1), which is valid for nonsymmetrical molecules. The triplet kind is called *orthohydrogen*. Its partition function is

$$z_{\text{rot}}^o = \sum_{\ell=1,3,5, \dots} (2\ell + 1) \exp[-\theta_{\text{rot}} \ell(\ell + 1)/T] \quad (27-8)$$

Since the multiplicity of the singlet state (parahydrogen) is 1 and that of the triplet (orthohydrogen) is 3, hydrogen gas behaves as though it were a mixture of one part,  $(1/4)N$ , of parahydrogen to three parts,  $(3/4)N$ , of orthohydrogen.

When the heat capacity of this mixture is measured in the usual way, the two kinds of hydrogen act like separate substances and the heat capacity  $C_V = T(\partial^2 F / \partial T^2)_V$  is a sum of two terms,

$$C_V^{\text{rot}} = NkT \left( \frac{1}{4} \frac{\partial^2}{\partial T^2} \ln z_{\text{rot}}^p + \frac{3}{4} \frac{\partial^2}{\partial T^2} \ln z_{\text{rot}}^o \right)_V \quad (27-9)$$

and not the single term

$$C_V^{\text{rot}} = NkT \left( \frac{\partial^2}{\partial T^2} \ln z_{\text{rot}} \right)_V \quad (27-10)$$

which was used in Chapter 22, was plotted in Fig. 22-1, and is valid for molecules with nonidentical nuclei. There is no difference between the two formulas in the classical limit of  $T > \theta_{\text{rot}}$ , where each partial is equal to unity; the result is  $NkT$  in both cases. However, at low temperatures there is a difference between (27-9) and (27-10), which is plotted in Fig. 27-1 [the dotted curve is for Eq. (27-10), the solid one for (27-9), which is the one predicted for  $H_2$ ]. The circles

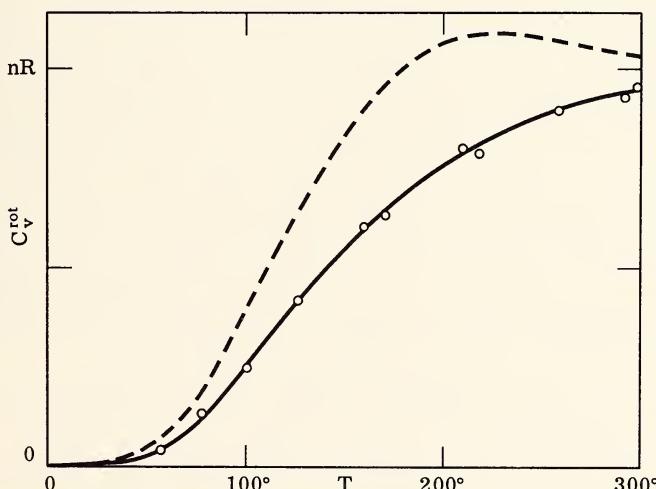


Fig. 27-1. The rotational part of the heat capacity of hydrogen gas ( $H_2$ ). It differs from the dashed curve (identical with Fig. 22-1) because the two nuclei are indistinguishable protons.

are the measured values, which definitely check with the assumption that para- and orthohydrogen act like separate substances.

If the heat capacities are measured very slowly, in the presence of a catalyst to speed the exchange of nuclear spin, then the two varieties of  $H_2$  do not behave like separate substances and the appropriate partition function is

$$z_{\text{rot}} = \frac{1}{4} \sum_{\ell \text{ even}} (2\ell + 1) \exp[-\theta_{\text{rot}} \ell(\ell + 1)/T] + \frac{3}{4} \sum_{\ell \text{ odd}} (2\ell + 1) \exp[-\theta_{\text{rot}} \ell(\ell + 1)/T] \quad (27-11)$$

and the heat capacity is obtained by inserting this in Eq. (27-10). This results in a different curve, which is not obtained experimentally without great difficulty. The metastable equilibrium, in which the two kinds of hydrogen behave as though they were different substances, is the usual experimental situation.

The nonsymmetric HD molecule, of course, has no ortho or para varieties; the two nuclei are not identical. The curve of Fig. 22-1, or the dotted curve of Fig. 27-1 (with a different  $\theta_{\text{rot}}$  of course) is the appropriate one here, and the one that corresponds to the measurements.

On the other hand the  $D_2$  molecule again has identical nuclei. The deuterium nucleus has one proton and one neutron and a spin of 1. Exchanging nuclei exchanges pairs of fermions, thus the wave function should be symmetric to exchange of nuclei. The molecule with antisymmetric spin factor (paradeuterium) is also antisymmetric in regard to spatial exchange ( $\ell = 1, 3, 5, \dots$ ) and the one with symmetric spin factor (orthodeuterium) is also symmetric in the spherical harmonic factor ( $\ell = 0, 2, 4, \dots$ ). There are twice as many ortho states as para states. Thus one can compute a still different curve for the  $C_V$  for  $D_2$  gas, which again checks with experiment. Because of the deep-lying requirements of symmetry of the wave function, the spin orientations of the inner nuclei, which can only be measured or influenced with the greatest of difficulty, reach out to affect the gross thermal behavior of the gas. These symmetry effects also correctly predict the factors  $\sigma$  in Eq. (22-4) for polyatomic molecules.

Thus the effects of quantum statistics turn up in odd corners of the field, at low temperatures and for substances a part of which can stay gaslike to low-enough temperatures for the effects of degeneracy to become evident. For the great majority of substances and over the majority of the range of temperature and density, classical statistical mechanics is valid, and the calculations using the canonical ensemble of Chapters 19 through 22 quite accurately portray the observed results. The situations where quantum statistics must be used to achieve concordance with experiment are in the minority (luckily; otherwise our computational difficulties would be much greater). But, when they are all considered, these exceptional situations add up to exhibit an impressive demonstration of the fundamental correctness of quantum statistics.

# References

The texts listed below have been found useful to the writer of this volume. They represent alternative approaches to various subjects treated here, or more complete discussions of the material.

E. Fermi, "Thermodynamics," Prentice-Hall, New York, 1937, is a short, readable discussion of the basic concepts.

W. P. Allis and M. A. Herlin, "Thermodynamics and Statistical Mechanics," McGraw-Hill, New York, 1952, presents some alternative approaches.

F. W. Sears, "Introduction to Thermodynamics, the Kinetic Theory of Gases, and Statistical Mechanics," Addison-Wesley, Reading, Mass., 1953, also provides some other points of view.

H. B. Callen, "Thermodynamics," Wiley, New York, 1960, is a "postulational" development of the subject.

Charles Kittel, "Elementary Statistical Physics," Wiley, New York, 1958, contains short dissertations on a number of aspects of thermodynamics and statistical mechanics.

J. C. Slater, "Introduction to Chemical Physics," McGraw-Hill, New York, 1939, has a more complete treatment of the application of statistical mechanics to physical chemistry.

L. D. Landau and E. M. Lifchitz, "Statistical Physics," Addison-Wesley, Reading, Mass., 1958, includes a thorough discussion of the quantum aspects of statistical mechanics.

# Problems

1. Suppose you are given a tank of gas that obeys the equation of state (3-4), a calibrated container that varies (slightly) in volume, in an unknown way, with temperature, and an accurate method of measuring pressure at any temperature. How would you devise a thermometer that measures the perfect gas scale of temperature? Could you also determine the constants  $a$  and  $b$  in the equation of state of the gas?
2. The coefficient of thermal expansion  $\beta$  and the compressibility  $\kappa$  of a substance are defined in terms of partial derivatives

$$\beta = (1/V)(\partial V / \partial T)_P \quad \kappa = -(1/V)(\partial V / \partial P)_T$$

- (a) Show that  $(\partial \beta / \partial P)_T = -(\partial \kappa / \partial T)_P$  and that  $(\beta / \kappa) = (\partial P / \partial T)_V$  for any substance.

- (b) It is found experimentally that, for a given gas,

$$\beta = \frac{RV^2(V - nb)}{RTV^3 - 2an(V - nb)^2} \quad \kappa = \frac{V^2(V - nb)^2}{nRTV^3 - 2an^2(V - nb)^2}$$

where  $a$  and  $b$  are constants, and also that the gas behaves like a perfect gas for large values of  $T$  and  $V$ . Find the equation of state of the gas.

3. A gas obeys equation of state (3-4). Show that for just one critical state, specified by the values  $T_c$  and  $V_c$ , both  $(\partial P / \partial V)_T$  and  $(\partial^2 P / \partial V^2)_T$  are zero. Write the equation of state giving  $P/P_c$  in terms of  $T/T_c$  and  $V/V_c$ . Plot three curves for  $P/P_c$  as a function of  $V/V_c$ , one for  $T = (1/2)T_c$ , one for  $T = T_c$ , and one for  $T = 2T_c$ . What happens physically when the equation indicates three allowed values of  $V$  for a single  $P$  and  $T$ ?

4. Suppose that all the atoms in a gas are moving with the same speed  $v$ , but that their directions of motion are at random.

- (a) Average over directions of incidence to compute the mean num-

ber of atoms striking an element of wall area  $dA$  per second (in terms of  $N$ ,  $V$ ,  $v$ , and  $dA$ ) and the mean momentum per second imparted to  $dA$ .

(b) Suppose, instead, that the number of atoms having speeds between  $v$  and  $v + dv$  is  $2N[1 - (v/v_m)](dv/v_m)$  for  $v < v_m$  (the directions still at random). Calculate for this case the mean number per second striking  $dA$  and the mean momentum imparted per second, in terms of  $N$ ,  $V$ ,  $v_m$ , and  $dA$ . Show that Eq. (2-4) holds for both of these cases.

5. A gas with Van der Waals' equation of state (3-4) has an internal energy

$$U = \frac{3}{2}nRT - (an^2/V) + U_0$$

Compute  $C_V$  and  $C_P$  as functions of  $V$  and  $T$  and compute  $T$  as a function of  $V$  for an adiabatic expansion.

6. An ideal gas for which  $C_V = (5/2)nR$  is taken from point  $a$  to point  $b$  in Fig. P-6, along three paths,  $acb$ ,  $adb$ , and  $ab$ , where  $P_2 = 2P_1$ ,  $V_2 = 2V_1$ .

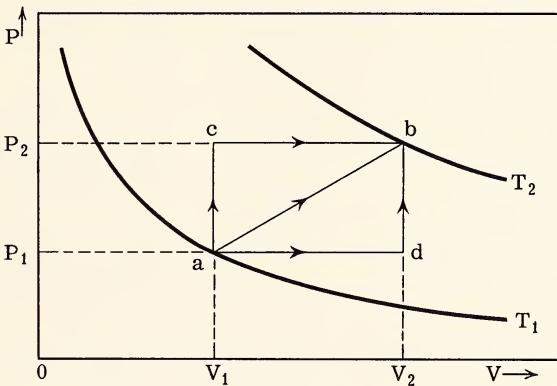


Fig. P-6

(a) Compute the heat supplied to the gas, in terms of  $N$ ,  $R$ , and  $T_1$ , in each of the three processes.

(b) What is the heat capacity of the gas, in terms of  $R$ , for the process  $ab$ ?

7. A paramagnetic solid, obeying Eqs. (3-6), (3-8), and (4-8) and having a heat capacity  $C_V = nAT^3$ , is magnetized isothermally (at constant volume) at temperature  $T_0$  from  $\mathfrak{M} = 0$  to a maximum magnetic field of  $\mathfrak{M}_m$ . How much heat must be lost? It is then demagnetized adiabatically (at constant volume) to  $\mathfrak{M} = 0$  again. Compute the final temperature  $T_1$  of the solid, in terms of  $\mathfrak{M}_m$ ,  $T_0$ ,  $A$ , and  $D$ . How do you explain away the fact that, if  $\mathfrak{M}_m$  is large enough or  $T_0$  small

enough, the formula you have obtained predicts that  $T_1$  should be negative?

8. Derive equations for  $(\partial V/\partial T)_P$  and  $(\partial V/\partial P)_T$  analogous to Eqs. (4-4) and (4-6). Obtain an expression for  $(\partial H/\partial P)_T$ . For a perfect gas, with  $C_p = (5/2)nRT$  and  $(C_p - C_v)/(\partial V/\partial T)_P = P$ , integrate the partials of  $H$  to obtain the enthalpy.

9. Figure P-9 shows a thermally isolated cylinder, divided into two parts by a thermally insulating, frictionless piston. Each side contains  $n$  moles of a perfect gas of point atoms. Initially both sides have the same temperature; heat is then supplied slowly and reversibly to the left side until its pressure has increased to  $(243P_0/32)$ .

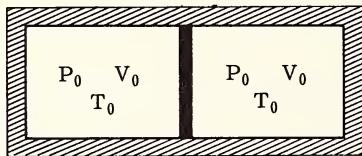


Fig. P-9

- How much work was done on the gas on the right?
- What is the final temperature on the right?
- What is the final temperature on the left?
- How much heat was supplied to the gas on the left?

10. The ratio  $C_p/C_v$  for air is equal to 1.4. Air is compressed from room temperature and pressure, in a diesel engine compression, to  $1/15$  of its original volume. Assuming the compression is adiabatic, what is the final temperature of the gas?

11. Compute  $\Delta Q_{12}$  and  $\Delta Q_{43}$ , for a Carnot's cycle using a perfect gas of point particles, in terms of  $nR$  and  $T_h$  and  $T_c$ . Using the perfect-gas scale of temperature, show that  $\Delta W_{23} = -\Delta W_{41}$ . Show that the efficiency of the cycle is  $(T_h - T_c)/T_h$  and thus prove that the perfect-gas temperature scale coincides with the thermodynamic scale of Eq. (5-5).

12. A magnetic material, satisfying Eqs. (4-8) and (3-8) has a constant heat capacity,  $C_{VM} = C$ . It is carried around a Carnot cycle shown in Fig. P-12,  $\mathfrak{M}$  being reduced isothermally from  $\mathfrak{M}_1$  to  $\mathfrak{M}_2$  at  $T_h$ , then reduced adiabatically from  $\mathfrak{M}_2$  to  $\mathfrak{M}_3$ , when it has temperature  $T_c$ , then remagnetized isothermally at  $T_c$  to  $\mathfrak{M}_4$ , and thence adiabatically back to  $T_h$  and  $\mathfrak{M}_1$ .

- Express  $\mathfrak{M}_3$  in terms of  $T_h$ ,  $T_c$ ,  $D$ ,  $C$ , and  $\mathfrak{M}_2$  and relate  $\mathfrak{M}_4$  similarly with  $\mathfrak{M}_1$ .
- How much heat is absorbed in process 1-2? How much given off in process 3-4?
- How much magnetic energy  $dW$  is given up by the material in each of the four processes? Show that  $dW_{23} = -dW_{41}$ .

(d) Show that the efficiency of the cycle heat-magnetic energy is  $(T_h - T_c)/T_h$ .

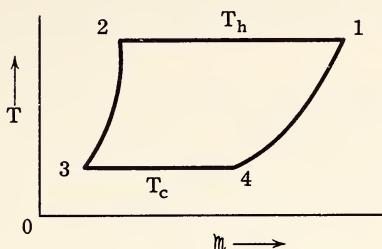


Fig. P-12

13. When a mole of liquid is evaporated at constant temperature  $T$  and vapor pressure  $P_v(T)$ , the heat absorbed in the process is called the latent heat  $L_v$  of vaporization. A Carnot cycle is run as shown in Fig. P-13, going isothermally from 1 to 2, evaporating  $n$  moles of liquid and changing volume from  $V_1$  to  $V_2$ , then cooling adiabatically to  $T - dT$ ,  $P_v - dP_v$  by evaporating an additional small amount of liquid, then recondensing the  $n$  moles at  $T - dT$ , from  $V_3$  to  $V_4$ , and thence adiabatically to  $P_v, T$  again.

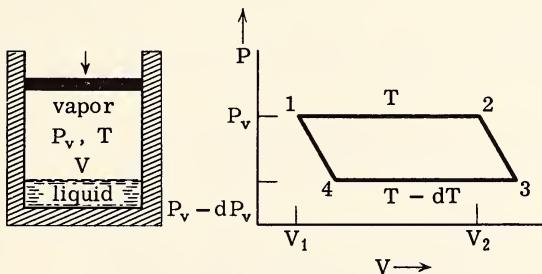


Fig. P-13

(a) Show that  $V_2 - V_1 = V_g - V_\ell$  where  $V_g$  is the volume occupied by  $n$  moles of the vapor and  $V_\ell$  the volume of  $n$  moles of the liquid and that if  $dT$  is small enough,  $V_3 - V_4 \approx V_2 - V_1$ .

(b) Find the efficiency of the cycle, in terms of  $dP_v$ ,  $V_g - V_\ell$ , and  $nL_v$ .

(c) If this cycle is to have the same efficiency as any Carnot cycle, this efficiency must be equal to  $(T_h - T_c)/T_h = dT/T$ . Equating the two expressions for efficiency, obtain an equation for the rate of change  $dP_v/dT$  of the vapor pressure with temperature in terms of  $V_g - V_\ell$ ,  $n$ ,  $L_v$ , and  $T$ .

14. Work out the Carnot cycle with a gas of photons, obeying Eqs. (7-8).

15. An ideal gas, satisfying Eqs. (4-7) and (4-12) is carried around the cycle shown in Fig. P-15; 1-2 is at constant volume, 2-3 is adiabatic, 3-1 is at constant pressure,  $V_3$  is  $8V_1$ , and  $n$  moles of the gas are used.

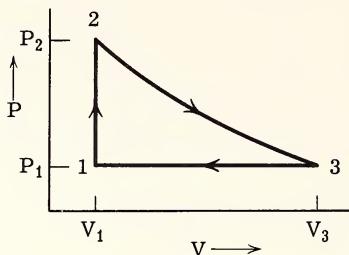
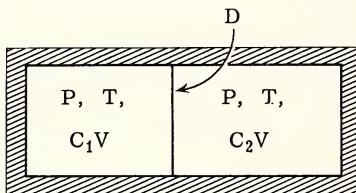


Fig. P-15

(a) What is the heat input, the heat output, and the efficiency of the cycle, in terms of  $P_1$ ,  $V_1$ ,  $n$ , and  $R$ ?

(b) Compare this efficiency with the efficiency of a Carnot cycle operating between the same extremes of temperature.

16. An amount of perfect gas of one kind is in the volume  $C_1V$  of Fig. P-16 at temperature  $T$  and pressure  $P$ , separated by an impervious diaphragm  $D$  from a perfect gas of another kind, in volume  $C_2V$  and at the same pressure and temperature ( $C_1 + C_2 = 1$ ). The volume  $V$  is isolated thermally. What is the entropy of the combination?



$$C_1 + C_2 = 1$$

Fig. P-16

Diaphragm  $D$  is then ruptured and the two gases mix spontaneously, ending at temperature  $T$ , partial pressure  $C_1P$  of the first gas,  $C_2P$  of the second gas, all in volume  $V$ . What is the entropy now? Devise a pair of processes, using semipermeable membranes (one of which will pass gas 1 but not 2, the other which will pass 2 but not 1), which will take the system from the initial to the final state reversibly and thus verify the change in entropy. What is the situation if gas 1 is the same as gas 2?

17. Two identical solids, each of heat capacity  $C_V$  (independent of  $T$ ), one at temperature  $T + t$ , the other at temperature  $T - t$ , may

be brought to common temperature  $T$  by two different processes:

(a) The two bodies are placed in thermal contact, insulated thermally from the rest of the universe and allowed to reach  $T$  spontaneously. What is the change of entropy of the bodies and of the universe caused by this process?

(b) First a reversible heat engine, with infinitesimal cycles, is operated between the two bodies, extracting work and eventually bringing the two to common temperature. Show that this common temperature is not  $T$ , but  $\sqrt{T^2 - t^2}$ . What is the work produced and what is the entropy change in this part of process b? Then heat is added reversibly to bring the temperature of the two bodies to  $T$ . What is the entropy change of the bodies during the whole of reversible process b? What is the change in entropy of the universe?

18. Show that  $T dS = C_V dT + T(\partial P/\partial T)_V dV$ , and  $T dS = C_V(\partial T/\partial P)_V dP + C_P(\partial T/\partial V)_P dV$ .

19. A paramagnetic solid, obeying Eqs. (3-6), (3-8), and (4-8), has a heat capacity  $C_{p0}(T)$  (at zero magnetic field) dependent solely on temperature. First, show that

$$T dS = C_P dT - T(\partial V/\partial T)_P dP + T(\partial \mathfrak{M}/\partial T)_{\mathfrak{M}} d\mathfrak{M}$$

and, analogous to Eq. (8-13), that  $(\partial C_{p0}/\partial \mathfrak{M})_{TP} = T(\partial^2 \mathfrak{M}/\partial T^2)_{\mathfrak{M}}$ . From this, show that

$$S = \int_0^T (C_{p0}/T) dT - \frac{1}{2} n D(\mathfrak{M}/T)^2 - \beta V_0 P$$

and thence obtain  $G$  and  $U$  as functions of  $T$ ,  $P$ , and  $\mathfrak{M}$ . Obtain  $S$  as a function of  $T$ ,  $V$ , and  $\mathfrak{M}$  and thence obtain  $F$  and  $U$  as functions of  $T$ ,  $V$ , and  $\mathfrak{M}$ .

20. A gas obeys the Van der Waals' equation of state (3-4) and has heat capacity at constant volume  $C_V = (3/2)nR$ . Write the equation of state in terms of the quantities  $t = T/T_c$ ,  $p = P/P_c$ , and  $v = V/V_c$ , where  $T_c = 8a/27Rb$ ,  $P_c = a/27b^2$ ,  $V_c = 3nb$  (see Problem 3). Calculate  $T_c S/P_c V_c$  in terms of  $t$  and  $v$ , likewise  $F/P_c V_c$  and  $G/P_c V_c$ . For  $t = 1/2$  plot  $p$  as a function of  $v$  from the equation of state. Then, for the same value of  $t$ , calculate and plot  $G/P_c V_c$  as a function of  $p$ , by graphically finding  $v$  for each value of  $p$  from the plot, and then computing  $G/P_c V_c$  for this value of  $v$  (remember that for some values of  $p$  there are three allowed values of  $v$ ). The curve for  $G/P_c V_c$  crosses itself. What is the physical significance of this?

21. Ten kilograms of water at a temperature of  $20^\circ\text{C}$  is converted to superheated steam at  $250^\circ$  and at the constant pressure of 1 atm. Compute the change of entropy of the water.  $C_p$  (water) =  $4180$  joules/kg-deg.  $L_v$  (at  $100^\circ\text{C}$ ) =  $22.6 \times 10^5$  joules/kg.  $C_p$  (steam) =  $1670 + 0.494T + 186 \times 10^6/T^2$  joules/kg-deg.

22. Assume that near the triple point the latent heats  $L_m$  and  $L_v$  are independent of  $P$ , that the vapor has the equation of state of a perfect gas, that the volume of a mole of solid or liquid is negligible compared to its vapor volume, and that the difference  $V_l - V_s$  is positive, independent of  $P$  or  $T$  and is small compared to  $nL_m/T$ . Using these assumptions, integrate the three Clausius-Clapeyron equations for the vapor-pressure, sublimation-pressure and melting-point curves. Sketch the form of these curves on the  $P$ - $V$  plane.

23. The heat of fusion of ice at its normal melting point is  $3.3 \times 10^5$  joules/kg and the specific volume of ice is greater than the specific volume of water at this point by  $9 \times 10^{-5}$  m<sup>3</sup>/kg. The value of  $(1/V)(\partial V/\partial T)_P$  for ice is  $16 \times 10^{-5}$  per degree and its value of  $-(1/V) \times (\partial V/\partial P)_T$  is  $12 \times 10^{-11}$  (m<sup>2</sup>/newton).

(a) Ice at  $-2^\circ\text{C}$  and at atmospheric pressure is compressed isothermally. Find the pressure at which the ice starts to melt.

(b) Ice at  $-2^\circ\text{C}$  and atmospheric pressure is kept in a container at constant volume and the temperature is gradually increased. Find the temperature at which the ice begins to melt.

24. Considering that all constituents of a chemical reaction are perfect gases obeying Eqs. (8-21), write out the expressions for  $\ln K$ , the logarithm of the equilibrium constant, in terms of  $T$ ,  $T_0$ , of the contents of integration  $g_{i0}$  and  $s_{i0}$  per mole and of the  $\nu_i$ 's. Show that the derivative of  $\ln K_2$  with respect to  $T$  at constant  $P$  is equal to  $\Delta H/RT^2$ , where  $\Delta H$  is the change in enthalpy (i.e., the heat evolved) when  $\nu_i$  moles of substance  $M_i$  disappears in the reaction (a negative value of  $\nu_i$  means the substance appears, i.e., is a product). This relation between  $[\partial \ln(K)/\partial T]_P$  and  $\Delta H$  is known as the van't Hoff equation.

25. The probability that a certain trial (throw of a die or drop of a bomb, for example) is a success is  $p$  for every trial. Show that the probability that  $m$  successes are achieved in  $n$  trials is

$$P_m(n) = \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m} \quad (\text{this is the binomial distribution})$$

Find the average number  $\bar{m}$  of successes in  $n$  trials, the mean-square ( $\bar{m}^2$ ) and the standard deviation  $\Delta m$  of successes in  $n$  trials.

26. The probability of finding  $n$  photons of a given frequency in an enclosure that is in thermal equilibrium with its walls is  $P_n = (1 - \alpha)\alpha^n$ , where  $\alpha$  ( $0 < \alpha < 1$ ) is a function of temperature, volume of the enclosure, and the frequency of the photons. What is the mean number  $\bar{n}$  of photons of this frequency? What is the fractional deviation  $\Delta n/\bar{n}$  of this number from the mean? Express this fractional deviation in terms of  $\bar{n}$ , the mean number. For what limiting value of  $\bar{n}$  does the fractional deviation tend to zero?

27. A molecule in a gas collides from time to time with another molecule. These collisions are at random in time, with an average interval  $\tau$ , the mean free time. Show that, starting at time  $t = 0$  (not an instant of collision) the probability that the molecule has not yet had its *next* collision at time  $t$  is  $e^{-t/\tau}$ . What is the expected time to this *next* collision? Show also that the probability that its *previous* collision (the last one it had before time  $t = 0$ ) was earlier than time  $-T$  is  $e^{-T/\tau}$ . What is the mean time of this *previous* collision? Does this mean that the average time interval between collisions is  $2\tau$ ? Explain the paradox.

28. In interstellar space, the preponderant material is atomic hydrogen, the mean density being about 1 hydrogen atom per cc. What is the probability of finding no atom in a given cc? Of finding 3 atoms? How many H atoms cross into a given cc, through one of its 1-cm<sup>2</sup> faces, per second, if the temperature is 1°K? If T is 1000°K?

29. A closed furnace F (Fig. P-29) in an evacuated chamber contains sodium vapor heated to 1000°K. What is the mean speed  $\bar{v}$  of the vapor atoms? At  $t = 0$  an aperture is opened in the wall of the furnace, allowing a collimated stream of atoms to shoot out into the vacuum.

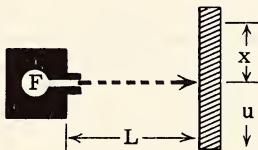


Fig. P-29

The aperture is closed again at  $t = \tau$ . A distance L from the aperture, a plate is moving with velocity  $u$ , perpendicular to the atom stream, so that the stream deposits its sodium atoms along a line on the plate; the position of the stream that strikes at time  $t$  hits the line at a point a distance  $X = ut$  from its beginning. Obtain a formula for the density of deposition of sodium as function of  $X$  along the line assuming that  $\tau \ll (L/\bar{v})$ , and find the value of  $X$  for which this density is maximum. Sketch a curve of the density versus  $X$ .

30. Most conduction electrons in a metal are kept from leaving the metal by a sudden rise in electric potential energy, at the surface of the metal, of an amount  $eW_0$ , where  $W_0$  is the electric potential difference between the inside and the outside of the metal. Show that if the conduction electrons inside the metal are assumed to have a Maxwell distribution of velocity, there will be a thermionic emission current of electrons from the surface of a metal at temperature T that is proportional to  $\sqrt{T} \exp(-eW_0/kT)$ . What is the velocity distribution of these electrons just outside the surface? [The measured

thermionic current is proportional to  $T^2 \exp(-e\phi/kT)$ , where  $\phi < W_0$ ; see Problem 63.]

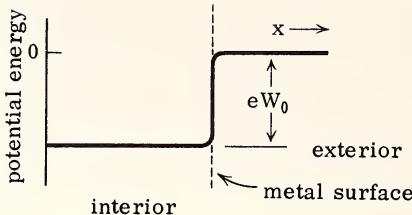


Fig. P-30

31. A gas of molecules with a Maxwell distribution of velocity at temperature  $T$  is in a container having a piston of area  $A$ , which is moving outward with a velocity  $u$  (small compared to  $\bar{v}$ ), expanding the gas adiabatically (Fig. P-31). Show that, because of the motion of the piston, each molecule that strikes the piston with velocity  $v$  at

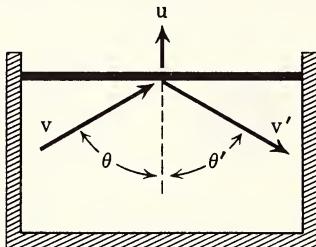


Fig. P-31

an angle of incidence  $\theta$  rebounds with a loss of kinetic energy of an amount  $2mvu \cos \theta$  ( $u \ll v$ ). Show that consequently the gas loses energy, per second, by an amount  $-dU = PAu = P dV$ , where  $dV$  is the increase in volume of the container per second.

32. Helium atoms have a collision cross section approximately equal to  $2 \times 10^{-16} \text{ cm}^2$ . In helium gas at standard conditions (1 atm pressure,  $0^\circ\text{C}$ ), assuming a Maxwell distribution, what is the mean speed of the atoms? What is their mean distance apart? What is the mean free path? The mean free time?

33. A gas is confined between two parallel plates, one moving with respect to the other, so that there is a flow shear in the gas, the mean gas velocity a distance  $y$  from the stationary plate being  $\beta y$  in the  $x$  direction (Fig. P-33). Show that the zero-order velocity distribution in the gas is

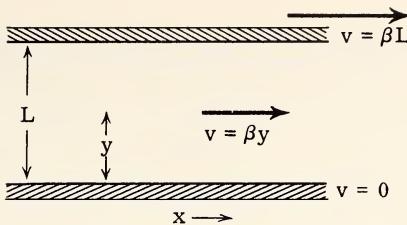


Fig. P-33

$$(1/2\pi mkT)^{3/2} \exp \{ -(1/2mkT)[(p_x - m\beta y)^2 + p_y^2 + p_z^2] \} = \\ = f_0(p_x - m\beta y, p_y, p_z)$$

Use Eq. (14-3) to compute  $f$  to the first order of approximation. Show that the mean rate of transport of  $x$  momentum across a unit area perpendicular to the  $y$  axis is  $(n/V)(t_c\beta/m) \iiint p_y^2 f_0 dp_x dp_y dp_z = (N/V)t_c\beta kT$ , which equals the viscous drag of the gas per unit area of the plate, which equals the gas viscosity  $\eta$  times  $\beta$ , the velocity gradient. Express  $\eta$  in terms of  $T$  and  $\lambda$  (mean free path) and show that the diffusion constant  $D$  of Eq. (14-8) is equal to  $(\eta/\rho)$ , where  $\rho$  is the density of the gas.

34. Use the Maxwell-Boltzmann distribution to show that, if the atmosphere is at uniform temperature, the density  $\rho$  and pressure  $P$  a distance  $z$  above the ground is  $\exp(-mgz/kT)$  times  $\rho_0$  and  $P_0$ , respectively (where  $g$  is the acceleration of gravity). Express  $\rho_0$  and  $P_0$  in terms of  $g$ ,  $T$ , and  $M_a$ , the total mass of gas above a unit ground area. Obtain this same expression from the perfect gas law,  $P = \rho kT/m$  and the equation  $dP = -\rho g dz$  giving the fall-off of pressure with height (assuming  $T$  is constant). Find the corresponding expressions for  $\rho$  and  $P$  in terms of  $z$  if the temperature varies with pressure in the way an adiabatic expansion does, i.e.,  $P = (\rho/C)^\gamma$ ,  $T = (D\rho)^{\gamma-1}$ , where  $\gamma = (C_p/C_v)$  [see Eqs. (4-12)].

35. A tube of length  $L = 2$  m and of cross section  $A = 10^{-4}$  m $^2$  contains CO<sub>2</sub> at normal conditions of pressure and temperature (under these conditions the diffusion constant  $D$  for CO<sub>2</sub> is about 10 m $^2$ /sec). Half the CO<sub>2</sub> contains radioactive carbon, initially at full concentration at the left-hand end, zero concentration at the right-hand end, the concentration varying linearly in between. What is the value of  $t_c$  for CO<sub>2</sub> under these conditions? Initially, how many radioactive molecules per second cross the mid-point cross section from left to right? [Use Eqs. (14-7).] How many cross from right to left? Compute the difference and show that it checks with the net flow, calculated from the diffusion equation (net flow) =  $-D(dn/dx)$ .

36. The collision cross section of an air molecule for an electron is about 10 $^{-19}$  m $^2$ . At what pressure will 90 per cent of the electrons emitted from a cathode reach an anode 20 cm away?

37. A gas-filled tube is whirled about one end with an angular velocity  $\omega$ . Find the expression for the equilibrium density of the gas as a function of the distance  $r$  from the end of the tube.

38. A vessel containing air at standard conditions is radiated with x-rays, so that 0.01 per cent of its molecules are ionized. A uniform electric field of  $10^4$  volts/meter is applied. What is the initial net flux of electrons? Of ions? (See Problem 36 for the cross section for electrons; the cross section for the ions is four times this. Why?) What is the ratio between drift velocity and mean thermal velocity for the electrons? For the ions?

39. A solid cylinder of mass  $M$  is suspended from its center by a fine elastic fiber so that its axis is vertical. A rotation of the cylinder through an angle  $\theta$  from equilibrium requires a torque  $K\theta$  to twist the fiber. When suspended in a gas-filled container at temperature  $T$ , the cylinder exhibits rotational fluctuation due to Brownian motion. What is the standard deviation ( $\Delta\theta$ ) of the amplitude of rotation and what is the standard deviation ( $\Delta\omega$ ) of its angular velocity? What would these values be if the container were evacuated?

40. Observations of the Brownian motion of a spherical particle of radius  $4 \times 10^{-7}$  m in water, at  $T = 300^\circ\text{K}$ , and of viscosity  $10^{-3}$  newton-sec/m<sup>2</sup> were made every 2 sec. The displacements in the  $x$  direction,  $x(t) - x(t-2)$ , were recorded and were tabulated as shown in Table P-40.

TABLE P-40

$\delta = x(t) - x(t-2)$ in units of $10^{-6}$ m	No. times this was observed
Less than $\pm 0.5$	111
Between 0.5 and 1.5	87
-0.5      -1.5	95
1.5      2.5	47
-1.5      -2.5	32
2.5      3.5	8
-2.5      -3.5	15
3.5      4.5	3
-3.5      -4.5	2
4.5      5.5	0
-4.5      -5.5	1
Larger than $\pm 5.5$	0

Compute the mean value of  $\delta$  and its standard deviation. How close is this distribution to a normal distribution [Eq. (11-17)]? Use Eq. (15-11) to compute Avogadro's number from the data, assuming  $R = kN_0$  is known.

41. Show that if the Hamiltonian energy of a molecule depends on a generalized coordinate  $q$  or momentum  $p$  in such a way that  $H \rightarrow \infty$  as  $p$  or  $q \rightarrow \pm \infty$ , it is possible to generalize the theorem on equipartition of energy to

$$\left. q \frac{\partial H}{\partial q} \right)_{av} = \left. \left( p \frac{\partial H}{\partial p} \right) \right)_{av} = kT$$

Verify that this reduces to ordinary equipartition when  $H$  has a quadratic dependence on  $q$  or  $p$ . If  $H$  has the relativistic dependence on the momentum

$$H = c \sqrt{(p_x^2 + p_y^2 + p_z^2) + m^2 c^2}$$

show that

$$(c^2 p_x^2 / H)_{av} = \dots = (c^2 p_z^2 / H)_{av} = kT$$

42. A harmonic oscillator has a Hamiltonian energy  $H$  related to its momentum  $p$  and displacement  $q$  by the equation

$$p^2 + (m\omega q)^2 = 2mH$$

When  $H = E$ , a constant energy, sketch the path of the system point in two-dimensional phase space. What volume of phase space does it enclose? In the case of  $N$  similar harmonic oscillators, which have the total energy  $E$  given by

$$\sum_{j=1}^N p_j^2 + \sum_{j=1}^N (M\omega q_j)^2 = 2ME$$

with additional coupling terms, too small to be included, but large enough to ensure equipartition of energy, what is the nature of the path traversed by the system point? Show that the volume of phase space "enclosed" by this path is  $(1/N!)(2\pi E/\omega)^N$ .

43. A gas of  $N$  point particles, with negligible (but not zero) collision interactions, enclosed in a container of volume  $V$ , has a total energy  $U$ . Show that the system point for the gas may be anywhere on a surface in phase space which encloses a volume  $V^{3N/2} (2\pi m U)^{3N/2} / (3N/2)!$ . For an ensemble of these systems to represent an equilibrium state, how must the system points of the ensemble be distributed over this surface?

44. A system consists of three distinguishable molecules at rest, each of which has a quantized magnetic moment, which can have its  $z$  component  $+M$ ,  $0$ , or  $-M$ . Show that there are 27 different possible

states of the system; list them all, giving the total  $z$  component  $M_{zi}$  of the magnetic moment for each. Compute the entropy  $S = -k \sum f_i \times \ln(f_i)$  of the system for the following a priori probabilities  $f_i$ :

(a) All 27 states equally likely (no knowledge of the state of the system).

(b) Each state is equally likely for which the  $z$  component  $M_z$  of the total magnetic moment is zero;  $f_i = 0$  for other states (we know that  $M_z = 0$ ).

(c) Each state is equally likely for which  $M_z = M$ ;  $f_i = 0$  for all other states (we know that  $M_z = M$ ).

(d) Each state is equally likely for which  $M_z = 3M$ ;  $f_i = 0$  for all other states (we know that  $M_z = 3M$ ).

(e) The distribution for which  $S$  is maximum, subject to the requirements that  $\sum f_i = 1$  and the mean component  $\sum f_i M_{zi}$  is equal to  $\gamma M$ . Show that for this distribution

$$f_i = \exp[(3M - M_{zi})\alpha / (1 + x + x^2)^3]$$

where  $x = e^\alpha M$  ( $\alpha$  being the Lagrange multiplier) and where the value of  $x$  (thus of  $\alpha$ ) is determined by the equation  $\gamma = 3(x^2 - 1)/(1 + x + x^2)$ . Compute  $x$  and  $S$  for  $\gamma = 3$ ,  $\gamma = 1$ , and  $\gamma = 0$ . Compare with a, b, c, and d.

45. Suppose the atoms of the crystal of Eqs. (20-6) are sufficiently "decoupled" so that it is a better approximation to consider the system as a collection of  $\nu = 3N$  harmonic oscillators, all having the same frequency  $\omega$ . Show that the partition function in this case is

$$Z = e^{-E_0/kT} / \left(1 - e^{-\hbar\omega/kT}\right)^{3N}$$

where

$$E_0 = [(V - V_0)^2/2\kappa V_0] + \frac{3}{2}N\hbar\omega$$

Compute the entropy, internal energy, and heat capacity of this system and obtain approximate formulas for  $kT$  small and also  $kT$  large compared to  $\hbar\omega$ . For what range of temperature do these formulas differ most markedly from those of Eq. (20-14)?

46. The atoms of the crystal of Problem 45 have spin magnetic moments, so that the allowed energy of interaction with the magnetic field of each atom is  $\mathcal{M}M_z = \pm(1/2)m\mathcal{M}$ , the magnets being parallel or anti-parallel respectively to the magnetic field  $\mathcal{M}$ . Show that for this system the canonical ensemble yields the following expression for the Helmholtz function:

$$F = E_0 + 3NkT \ln \left(1 - e^{-\hbar\omega/kT}\right) - \frac{1}{2}Nm\mathcal{M} - NkT \ln \left(1 + e^{-m\mathcal{M}/kT}\right)$$

$$\rightarrow E_0 + 3NkT \ln(\hbar\omega/kT) - NkT \ln 2 - (Nm\mathcal{K}^2/8kT)$$

$$kT \gg \hbar\omega \text{ and } m\mathcal{K}$$

Compute  $S$ ,  $C_V$ , and  $U$  for the high-temperature limit and compare the results with Problem 19. What is the magnetization  $m\mathcal{M}$  for this system? What is the rate of change of  $T$  with respect to  $\mathcal{K}$  for adiabatic demagnetization of this crystal at constant volume?

47. Use the final result of Problem 42 to show that the entropy of  $N$  distinguishable harmonic oscillators, according to the microcanonical ensemble (every system in the ensemble has energy  $NkT$ ) is

$$S = -Nk[1 + \ln(kT/\hbar\omega)]$$

48. For the solid described by Eq. (20-16) show that  $P = [(V_0 - V)/\kappa V_0] + (\gamma U_D/V)$ , where  $U = [(V_0 - V)^2/2\kappa V_0] + U_D$  and  $\gamma = (-V/\theta) \times (d\theta/dV)$ . Thence show that, for any temperature, if  $\gamma$  is independent of temperature, the thermal expansion coefficient  $\beta$  is related to  $\gamma$  by the formula

$$\beta = (1/V)(\partial V/\partial T)_P = \kappa(\partial P/\partial T)_V = (\kappa\gamma C_V/V)$$

Constant  $\gamma$  is called the Grüneisen constant.

49. A system consists of a box of volume  $V$  and a variable number of indistinguishable (MB) particles each of mass  $m$ . Each particle can be "created" by the expenditure of energy  $\gamma$ ; once created it becomes a member of a perfect gas of point particles within the volume  $V$ . The allowed energies of the system are therefore  $n\gamma$  plus the kinetic energies of  $n$  particles inside  $V$ , for  $n = 0, 1, 2, \dots$ . Show that the Helmholtz function for this system (canonical ensemble) is

$$F = kT \ln \left[ \sum_{n=0}^{\infty} \left( V^n \bar{X}^n / n! \right) \right] = -kTV\bar{X}$$

where  $\bar{X} = (2\pi mkT/h^2)^{3/2} e^{-\gamma/kT}$ . Calculate the probability that  $n$  particles are present in the box and thence obtain an expression for  $\bar{N}$ , the mean number of free particles present as a function of  $\gamma$ ,  $T$ , and  $V$ . Also calculate  $S$ ,  $C_V$ , and  $P$  from  $F$  and express these quantities as functions of  $\bar{N}$ ,  $T$ , and  $V$ .

50. What fraction of the molecules of  $H_2$  gas are in the first excited rotational state ( $\ell = 1$ ) at  $20^\circ K$ , at  $100^\circ K$ , and at  $5000^\circ K$ ? What are the corresponding fractions for  $O_2$  gas? What fraction of the molecules of  $H_2$  gas are in the first excited vibrational states ( $n = 1$ ) at  $20^\circ K$  and  $5000^\circ K$ ? What are the corresponding fractions for  $O_2$  gas?

51. Plot the heat capacity  $C_V$ , in units of  $Nk$  for  $O_2$  gas, from 100 to  $5000^\circ K$ .

52. The solid of Eqs. (20-14) sublimes at low pressure, at a sublimation temperature  $T_S$  which is large compared to  $\theta$ , the resulting vapor being a perfect diatomic gas, with properties given by Eqs. (21-14) and (22-2) (where  $\theta_{\text{rot}} \ll T_S \ll \theta_{\text{vib}}$ ). Show that the equation relating  $T_S$  and the sublimation pressure  $P_S$  is approximately

$$G_S \simeq V_0 P_S + \frac{3}{2} Nk\theta + 3NkT_S \ln(\theta/T_S) - NkT_S$$

$$= G_g \simeq NkT_S \ln(P_S V_0 T_S^{7/2} / N_0 k \theta^{5/2}) - NkT_S$$

where the equation

$$N_0 = V_0 (4\pi I e k \theta / h^2) (2\pi m k \theta / h^2)^{3/2}$$

defines the constant  $N_0$ . Since  $V_0 \ll V_g = NkT_S/P_S$  and  $\theta \ll T_S$ , show that this reduces to

$$P_S \simeq N_0 k \sqrt{T\theta} / V_0$$

Also show that the latent heat of sublimation

$$L_S = T_S (S_g - S_S) \simeq \frac{1}{2} NkT_S$$

53. Work out the grand canonical ensemble for a gas of point atoms, each with spin magnetic moment, which can have magnetic energy  $+(1/2)\mu\mathcal{C}$  or  $-(1/2)\mu\mathcal{C}$  in a magnetic field  $\mathcal{C}$  in addition to its kinetic energy. Obtain the expression for  $\bar{N}$  and expressions for  $\Omega$ ,  $\mu$ ,  $S$ ,  $U$ ,  $C_V$ , and the equation of state, in terms of  $\bar{N}$ ,  $T$ , and  $\mathcal{C}$ . How much heat is given off by the gas when the magnetic field is reduced from  $\mathcal{C}$  to zero isothermally, at constant volume?

54. A system consists of three particles, each of which has three possible quantum states, with energy 0,  $2E$ , or  $5E$ , respectively. Write out the complete expression for the partition function  $Z$  for this system: (a) if the particles are distinguishable; (b) if the particles obey Maxwell-Boltzmann statistics; (c) if they obey Einstein-Bose statistics; (d) if they obey Fermi-Dirac statistics. Calculate the entropy of the system in each of these cases.

55. The maximum intensity per unit frequency interval, in the sun's spectrum, occurs at a wavelength of  $5000 \text{ \AA}$ . What is the surface temperature of the sun?

56. Show that, for Einstein-Bose particles (bosons)

$$S = -k \sum_i [\bar{n}_i \ln(\bar{n}_i) - (1 + \bar{n}_i) \ln(1 + \bar{n}_i)]$$

57. It has been reported that the fission bomb produces a temperature of a million °K. Assuming this to be true over a sphere 10 cm in diameter: (a) What is the radiant-energy density inside the sphere? (b) What is the rate of radiation from the surface? (c) What is the radiant flux density 1 km away? (d) What is the wavelength of maximum energy per unit frequency interval?

58. The Planck distribution can be obtained by considering each standing electromagnetic wave in a rectangular enclosure ( $L_x L_y L_z$ ) as a degree of freedom, with coordinate  $Q_\nu$  proportional to the amplitude of the electric vector, with momentum  $P_\nu$  proportional to the amplitude of the magnetic vector and with a field energy, corresponding to a term in the Hamiltonian, equal to  $2\pi c^2 P_\nu^2 + (\omega_\nu^2/8\pi c^2) Q_\nu^2$ , where  $c$  is the velocity of light and where the allowed frequency of the  $\nu$ -th standing wave is given by

$$\omega_\nu^2 = \pi^2 c^2 [(k_\nu/L_x)^2 + (m_\nu/L_y)^2 + (n_\nu/L_z)^2]$$

(because of polarization, there are two different waves for each trio  $k_\nu, m_\nu, n_\nu$ ). Use the methods of Eqs. (20-4) to (20-11) to prove that the average energy contained in those standing waves with frequencies between  $\omega$  and  $\omega + d\omega$  is  $dE = (\hbar/\pi^2 c^3) \omega^3 d\omega / (e^{\hbar\omega/kT} - 1)$ . Compare this derivation with the one dealing with photons, which produced Eq. (25-3).

59. Analyze the thermal oscillations of electromagnetic waves along a conducting wire of length  $L$ . In this case of one-dimensional, standing waves, the  $n$ -th wave will have the form  $\cos(\pi n x/L) e^{i\omega t}$ , where  $\omega = 2\pi f = \pi n c / L$ ,  $c$  being the wave velocity and  $f$  the frequency of the  $n$ -th standing wave. Use a one-dimensional analogue of the derivation of Problem 58 to show that the energy content of the waves with frequencies between  $f$  and  $f + df$  is  $[2Lhf df/c(e^{hf/kT} - 1)]$ . If the wire is part of a transmission line, which is terminated by its characteristic impedance, all this energy will be delivered to this impedance in a time  $2L/C$ . Show, consequently, that the power delivered to the terminal impedance, the thermal noise power, in the frequency band  $df$  at frequency  $f$  is  $hf df / (e^{hf/kT} - 1)$ . Show that this reduces to the familiar uniform distribution (white noise) at low frequencies, but that at high frequencies the power drops off exponentially. Below what frequency is the noise power substantially "white" at room temperatures (300°K)?

60. A container of volume  $V$  has  $N$  short-range attractive centers (potential wells) fixed in position within the container. There are also bosons within the container. Each particle can either be bound to an

attractive center, with an energy  $-\gamma$  (one level per center), or can be a free boson, with energy equal to its kinetic energy,  $E$ . Use the analysis of Chapter 30 to show that the equation relating the mean number  $\bar{N}$  of bosons to their chemical potential  $\mu$  is

$$\bar{N} = \frac{N}{e^{(\gamma - \mu)/kT} - 1} + 1.129N_0 f_{1/2}\left(\frac{-\mu}{kT}\right); \quad N_0 = gV \left(\frac{2\pi mkT}{h^2}\right)^{3/2}$$

Draw curves for  $-\mu/kT$  as a function of  $N_0/N$  for  $\bar{N}/N = 1$  and for  $\gamma/kT = 0.1$  and  $1.0$ , using Table 25.1. Draw the corresponding curves for  $PV/NkT$ .

61. Suppose the particles of Problem 60 are MB particles instead of bosons. Calculate the partition function  $Z$  for a canonical ensemble and compare it with the  $Z$  for Problem 49.

62. Show that, for Fermi-Dirac particles (fermions),

$$S = - \sum_i [\bar{n}_i \ln(\bar{n}_i) + (1 - \bar{n}_i) \ln(1 - \bar{n}_i)]$$

63. The conduction electrons of Problem 30 are, of course, fermions. Show that, for FD statistics, the thermionic emission current from the metal surface at temperature  $T$  is proportional to  $T^2 \exp(-e\phi/kT)$ , where  $\phi = W_0 - \mu \approx W_0 - \mu_0$  is called the thermionic work function of the surface.

64. The container and  $N$  attractive centers of Problem 60 have  $N$  fermions, instead of bosons, in the system. By using Eqs. (26-7) and (26-8) show that the equation relating  $\mu$  and  $T$  and  $V$  is

$$\bar{N} = \frac{N}{e^{(\gamma - \mu)/kT} + 1} + N_0 \eta(\mu/kT), \quad N_0 \text{ as in Problem 60}$$

Plot  $\mu/kT$  as function of  $N_0/N$  for  $\gamma/kT = 0.1$  and  $1.0$ , using Table 26-1. Draw the corresponding curves for  $PV/NkT$ .

65. Calculate the heat capacity of  $D_2$  as a function of  $T/\theta_{\text{rot}}$  from 0 to 1.

66. The Schrödinger equation for a one-dimensional harmonic oscillator is

$$H\psi \equiv -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + \frac{1}{2}m\omega^2 x^2 \psi = E\psi$$

Its allowed energies and corresponding wave functions are

$$E_n = \hbar\omega \left(n + \frac{1}{2}\right)$$

$$\psi_n(x) = \sqrt{m\omega/2^n n! \hbar \sqrt{\pi}} \quad H_n(x \sqrt{m\omega/\hbar}) \exp(-m\omega x^2/2\hbar)$$

where  $H_0(z) = 1$ ,  $H_1(z) = 2z$ ,  $H_2(z) = 4z^2 - 2$ ,  $H_3(z) = 8z^3 - 12z$ , etc.

Two identical, one-dimensional oscillators thus have a Schrödinger equation

$$H(x,y)\Psi \equiv -\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \Psi + \frac{1}{2} m\omega^2(x^2 + y^2) \Psi = E\Psi$$

where  $x$  is the displacement of the first particle from equilibrium and  $y$  that of the second.

(a) Show that allowed solutions of this equation, for the energy  $\hbar\omega(n + 1/2)$ , may be written either as linear combinations of the products  $\psi_m(x)\psi_{n-m}(y)$  for different values of  $m$  between 0 and  $n$ , or else as linear combinations of the products

$$\psi_m\left(\frac{x+y}{\sqrt{2}}\right) \psi_{n-m}\left(\frac{x-y}{\sqrt{2}}\right)$$

(b) Express the solutions

$$\psi_m\left(\frac{x+y}{\sqrt{2}}\right) \psi_0\left(\frac{x-y}{\sqrt{2}}\right), \quad \psi_m\left(\frac{x+y}{\sqrt{2}}\right) \psi_1\left(\frac{x-y}{\sqrt{2}}\right),$$

and

$$\psi_m\left(\frac{x+y}{\sqrt{2}}\right) \psi_2\left(\frac{x-y}{\sqrt{2}}\right)$$

for  $m = 0, 1$ , and  $2$  as linear combinations of the solutions  $\psi_m(x)\psi_n(y)$  for  $m, n = 0, 1, 2$ .

(c) Which of these solutions are appropriate if the two particles are bosons? Which if they are fermions?

(d) Suppose the potential energy has an interparticle repulsive term  $-(1/2)m\kappa^2(x - y)^2$  (where  $\kappa^2 < \omega^2$ ) in addition to the term  $(1/2)m\omega^2(x^2 + y^2)$ . Show that, in this case, the allowed energies for bosons differ from those for fermions. Which lie higher and why?

# Constants

Gas constant  $R = 1.988 \text{ kg-cal/mole-deg.}$

$= 8.314 \times 10^3 \text{ joules/mole-deg.}$

Avogadro's number  $N_0$  (No. per kg mole)  $= 6.025 \times 10^{26}$

Number of molecules per  $\text{m}^3$  at standard conditions  $= 2.689 \times 10^{25}$

Standard conditions:  $T = 0^\circ\text{C} = 273.16^\circ\text{K}$ ,  $P = 1 \text{ atm}$

Volume of perfect gas at standard conditions  $= 22.41 \text{ m}^3/\text{mole}$

Dielectric constant of vacuum  $\epsilon_0 = (1/9) \times 10^{-9} \text{ farads/m}$

Electronic charge  $e = 1.602 \times 10^{-19} \text{ coulomb}$

Electron mass  $m = 9.106 \times 10^{-31} \text{ kg}$

Proton mass  $= 1.672 \times 10^{-27} \text{ kg}$

Planck's constant  $h = 6.624 \times 10^{-34} \text{ joule-sec}$

$\hbar = (h/2\pi) = 1.054 \times 10^{-34}$

Boltzmann's constant  $k = 1.380 \times 10^{-23} \text{ joule/}^\circ\text{K}$

Acceleration of gravity  $g = 9.8 \text{ m/sec}^2$

1 atm  $= 1.013 \times 10^5 \text{ newtons/m}^2$

1 cm Hg  $= 1333 \text{ newtons/m}^2$

1 newton  $= 10^5 \text{ dynes}$

1 joule  $= 10^7 \text{ ergs}$

1 electronvolt (ev)  $= 1.59 \times 10^{-19} \text{ joules}$

$= k(7500^\circ\text{K})$

Velocity of 1 ev

electron  $= 5.9 \times 10^5 \text{ m/sec}$

1 kg-cal  $= 4182 \text{ joules}$

Ratio of proton

mass to elec-

tron mass  $= 1836$

$\bar{v} = \sqrt{8kT/m} = 1.96 \times 10^4 \text{ m/sec}$  for electrons at  $T = 1^\circ\text{K}$

$= 146 \text{ m/sec}$  for protons at  $T = 1^\circ\text{K}$

$h^2 N_0^{2/3} / 2\pi m k = 3961^\circ\text{K} - \text{m}^2$  for electrons,  $= 2.157^\circ\text{K} - \text{m}^2$  for protons

$h^3 N_0 / (2\pi m k T)^{3/2} = V_b / T^{3/2}$ ;  $V_b = \begin{cases} 2.49 \times 10^5 \text{ m}^3 & \text{for electrons} \\ 3.17 \text{ m}^3 & \text{for protons} \end{cases}$  [see Eq. (24-11)]

$$(\mu_0/kT)(VN_0/\bar{N})^{2/3} = (h^2/2mkT)(3N_0/8\pi)^{2/3} = A_d/T$$

$A_d = 3017 \text{ m}^2$  for electrons [see Eq. (31-3)]

# Glossary

Symbols used in several chapters are listed here with the numbers of the pages on which they are defined.

a	Van der Waals' constant, 18, 193	H	enthalpy, 52, 60
a	Stefan's constant, 55, 225	H	Hamiltonian function, 109, 141, 239
A	area, 10	$\mathcal{H}$	magnetic intensity, 15
b	Van der Waals' constant, 18, 193	I	amount of information, 195
$\mathfrak{B}$	magnetic induction, 15	J	torsion in bar, 15
c	specific heat, 15	k	Boltzman constant, 17, 147
C	heat capacity, 15	L	latent heat, 70
d	perfect differential, 21	m	mass of particle, 10
$\mathfrak{d}$	imperfect differential, 22	M	mobility, 119
D	Curie constant, 20, 115	$\mathfrak{M}$	magnetization, 15
D	diffusion constant, 121, 133	n	number of moles, 14
D(x)	Debye function, 178	$n_j$	occupation number, 298
e	= 2.7183, nat. log. base	N	number of particles, 10, 209
e	charge on electron, 118	$N_0$	Avogadro's number, 13
E	energy of system, 145	p	momentum, 96, 109, 141
$\mathfrak{E}$	potential difference, 83	$p_i$	partial pressure, 80
$\mathfrak{E}$	electric intensity, 118	P	pressure, 9, 165
f	distribution function, 91, 112, 141	$\mathfrak{P}$	magnetic polarization, 15
F	Helmholtz' function, 62, 165	q	coordinate, 108, 141
$\mathfrak{F}$	Faraday constant, 83	Q	heat, 8
g	multiplicity of state, 167, 227	Q	collision function, 104, 116
G	Gibbs' function, 63	r	internuclear distance, 190
h	Planck's constant, 55	$\mathbf{r}$	position vector, 103
$\mathfrak{h}$	= $h/2\pi$	R	gas constant, 16
$h_i$	scale factor, 108	S	entropy, 40, 45, 147, 165
		t	time, 104
		$t_c$	relaxation time, 116
		T	temperature, 7, 17, 38

U	internal energy, 9, 25, 59	$\nu$	number of degrees of freedom, 108
U	drift velocity, 119	$\nu$	quantum number, 145, 208
v	velocity, 10	$\nu_i$	stoichiometric coefficient, 78
V	volume, 10	$\pi$	= 3.1416
W	work, 23	$\rho$	density, 128, 235
x,y,z	coordinates, 103	$\sigma$	standard deviation, 90
Z	normalizing constant, 110	$\sigma$	collision cross section, 101
Z	partition function, 165, 182	$\tau$	mean free time, 102
$\mathcal{Z}$	grand partition function, 204	$\phi(q)$	potential energy, 105
$\alpha$	Lagrange multiplier, 151, 164	$\phi$	number of degrees of freedom, 141
$\beta$	thermal expansion coefficient, 19	$\chi$	magnetic susceptibility, 15
$\gamma$	$C_p/C_v$ , 19	$\chi_i$	concentration, 80
$\epsilon$	particle energy, 181, 208	$\Psi$	wave function, 238
$\eta$	heat efficiency, 34	$\omega$	$2\pi$ (frequency), 55, 170
$\eta$	viscosity, 131	$\omega_j$	oscillator constant, 124, 175
$\theta$	Debye temperature, 178	$\Omega$	grand potential, 63, 204
$\kappa$	compressibility, 19	$\approx$	approximately equal to
$\lambda$	mean free path, 102	$\langle \rangle$	average value, 89
$\mu$	chemical potential, 16, 45, 79, 205	$\partial$	partial derivative, 20
$\mu$	magnetic moment, 113	$\ln$	natural logarithm, 44
$\mu_0$	permeability of vacuum, 15	$n!$	factorial function, 94, 156, 159

# Index

Absolute zero, 17  
Adiabatic process, 29, 34, 45, 230, 251, 259  
Antisymmetric wave functions, 241, 267  
Average energy, 110, 164, 203  
kinetic energy, 10, 123  
value, 89  
Avogadro's number, 13, 260, 268  
Binomial distribution, 89, 256  
Black-body radiation, 54, 221, 265  
Boltzmann constant, 17, 147, 268  
equation, 104, 116  
Maxwell- (see Maxwell-Boltzmann distribution)  
Bose-Einstein statistics, 212, 264  
boson, 212, 220, 241, 266  
Calculations, thermodynamic, 64  
Canonical ensemble, 163, 202, 262  
Capacity, heat, 14, 113, 172, 197, 201, 231  
at constant magnetization, 29, 56, 255  
at constant pressure, 15, 28, 65, 251  
at constant volume, 15, 28, 65, 113, 236, 247, 255  
Carnot cycle, 33, 40, 252  
Centigrade temperature scale, 7, 16  
Chemical potential, 16, 45, 56, 67, 79, 203, 211, 234  
Chemical reactions, 78  
Clausius' principle, 37  
Clausius-Clapeyron equation, 71, 256  
Collision, 97, 101  
Collision function, 104, 116  
Compressibility, 19, 112, 155, 179  
Concentration, 80  
Condensation, Einstein, 228, 232  
Conditional probability, 68, 131  
Conductivity, electric, 118, 260  
Continuity, equation of, 104, 142  
Critical point, 74  
Cross section, collision, 101, 259  
Crystal, 111, 154, 170, 262  
Curie's law, 20, 56, 113  
Cycle  
    Carnot, 33, 40  
    reversible, 34  
Debye formulas for  $C_V$ , 177, 234  
Debye function, 178  
Degenerate state, 216, 230, 234  
Degrees centigrade, 7, 16  
Degrees of freedom, 108, 222  
Degrees Kelvin, 17, 38  
Density  
    fluctuations, 127, 206

Density, probability, 91  
 Dependent variables, 6  
 Derivative, partial, 20, 65  
 Deuterium, 248  
 Deviation, standard, 90, 98  
 Diatomic molecule, 195  
 Differential  
     imperfect, 22  
     perfect, 21  
 Diffusion, 120, 133, 259  
 Dirac (*see* Fermi-Dirac statistics)  
 Distinguishable particles, 186, 210  
 Distribution function, 89, 112, 141, 147, 216  
     binomial, 89, 256  
     geometric, 218, 256  
     Maxwell, 98, 160, 258  
     Maxwell-Boltzmann, 103, 105, 217, 233, 259  
     momentum, 94  
     normal, 93, 98  
     Planck, 222, 265  
     Poisson, 91, 206, 218  
 Drift velocity, 98, 119, 260

Efficiency of cycle, 34  
 Einstein (*see* Bose-Einstein statistics)  
     condensation, 231  
     formula for  $C_V$ , 171  
 Electric conduction in gases, 118, 260  
 Electrochemical processes, 82  
 Electromagnetic waves, 54, 221, 265  
 Electrons, 213, 234, 244, 266  
 Energy, 9  
     density, 54, 222  
     distribution in, 103, 165, 217, 233  
     Hamiltonian, 109, 141, 239, 261  
     internal, 9, 12, 23, 59, 155, 164, 204, 251  
     kinetic, 10, 108, 183  
     magnetic, 16, 125

Energy, potential, 105, 123  
     rotational, 195  
     vibrational, 112, 155, 200  
 Engine, heat, 33  
 Ensemble, 141, 147, 154, 163  
     canonical, 163, 202, 262  
     grand canonical, 202, 264  
     microcanonical, 154, 261  
 Enthalpy, 52, 60  
 Entropy, 40, 69, 76, 147, 165, 203  
     of mixing, 49, 254  
     of perfect gas, 45  
     of universe, 44, 255  
 Equation of state, 16, 18, 64, 165, 179, 193, 204, 250

Equilibrium constant, 81  
     state, 6, 59  
 Equipartition of energy, 110, 122  
 Euler's equation, 42, 64  
 Evaporation, 72  
 Expansion coefficient, thermal, 19, 250  
 Expected value, 89  
 Extensive variables, 14, 41, 59

Factorial function, 94, 156, 159  
 Fermi-Dirac statistics, 213, 233, 264  
 Fermion, 212, 233, 242, 266  
 Fluctuations, 122, 206, 216, 226, 260, 265  
 Fokker-Planck equation, 133

Gas  
     boson, 220  
     constant, R, 16, 268  
     fermion, 233  
     fluctuations in, 127, 206  
     paramagnetic, 56  
     perfect, 10, 17, 30, 45, 51, 157, 205, 252  
     photon, 54, 221, 265  
     statistical mechanics of, 181  
     Van der Waals, 18, 48, 193, 251, 255

Geometric distribution, 218, 256  
 Gibbs' function, 63, 70, 77, 203

Gibbs-Helmholtz equation, 83  
 Grand canonical ensemble, 202, 264  
 Grand partition function, 203, 209  
 Grand potential, 63, 204, 211, 213  
 Grüneisen constant, 263  
 Hamiltonian function, 109, 141, 239, 261  
 Hamilton's equations, 142, 173  
 Heat, 8, 24  
     capacity, 14, 28, 56, 65, 113, 231, 236, 247  
     engine, 33  
     reservoir, 26, 29, 34  
     specific, 15, 172, 197, 201, 229  
 Heisenberg principle, 145  
 Helium, 212, 232, 244  
 Helmholtz function, 62, 165, 178, 189, 204  
 Hydrogen gas, 196, 200, 245  
 Hypersphere, volume of, 159  
 Independent variables, 6  
 Indistinguishability, 185  
 Information, 147  
 Integrating factor, 22, 40  
 Intensive variables, 14, 41  
 Interaction between particles, 190  
 Internal energy, 9, 12, 23, 59, 155, 164, 204, 251  
 Inversion point, 52  
 Irreversible process, 42  
 Isothermal process, 29, 33  
 Joule coefficient, 47  
     experiment, 43, 46  
 Joule-Thomson coefficient, 51  
     experiment, 51  
 Kelvin, degrees, 17, 38  
 Kelvin's principle, 35  
 Kinetic energy, 10, 108, 183  
 Kinetic energy, per degree of freedom, 110, 122  
 Lagrange multipliers, 150, 164, 203  
 Langevin equation, 131  
 Latent heat, 70  
 Law of mass action, 81  
 Law of thermodynamics  
     first, 25  
     second, 32, 38  
 Legendre transformation, 61  
 Liouville's theorem, 131  
 Macrostate, 141  
 Magnetic induction, 15  
 Magnetic intensity, 15, 20  
 Magnetization, 15, 20, 30, 113, 263  
 Mass, 10  
 Mass action, law of, 81  
 Maxwell distribution, 90, 160, 258  
 Maxwell's relations, 60, 62  
 Maxwell-Boltzmann  
     distribution, 102, 165, 217, 233, 259, 264  
     particles, 209, 211, 217, 231  
 Mean free path, 102  
     time, 102, 257  
 Melting, 69  
     latent heat of, 70  
 Metals, 180, 234, 244, 266  
 Metastable equilibrium, 243  
 Microcanonical ensemble, 154, 261  
 Microstate, 5, 141, 146  
 Mixing, entropy of, 49, 254  
 Mobility, 119, 121  
 Moment, magnetic, 113, 262  
 Momentum, 11, 55, 142, 173  
     distribution (see Maxwell-Boltzmann distribution)  
     of photon, 55  
     space, 96  
 Multiplicity, 167, 185, 227, 233

Natural variables, 41, 57  
 Newton, 9, 268  
 Noise, thermal, 127, 206, 265  
 Normal distribution, 93, 98  
 Normal modes of crystal, 175  
 Normalizing constant, 110, 238  
 Numerator, bringing to, 65

Occupation numbers, 208, 218, 220  
 Ohm's law, 119  
 Omega space, 176  
 Orthohydrogen, 245  
 Orthogonal functions, 238  
 Oscillator, simple, 111, 154, 170, 261  
 fluctuations of, 125  
 Oxygen gas, 102, 196, 200

Parahydrogen, 245  
 Paramagnetic substance, 16, 20, 30, 56, 113, 251, 255  
 Partial derivative, 20, 65  
 Particle states, 185, 208  
 Partition function, 165, 182, 203, 209  
 Path, mean free, 102  
 Pauli principle, 213, 218, 242  
 Perfect differential, 21  
 Perfect gas, 17, 51, 96, 152, 157, 205  
 of point particles, 10, 30, 45  
 Phase, change of, 68  
 Phase space, 103, 141, 144  
 Phonons, 174  
 Photon gas, 54, 212, 221, 265  
 Planck distribution, 222, 265  
 Planck's constant, 55, 222, 268  
 Poisson distribution, 91, 206, 218  
 Polarization, magnetic, 15, 114  
 Potential  
 chemical, 16, 45, 56, 67, 79, 203, 211, 234  
 energy, 105, 123, 191, 199  
 grand, 63, 204, 211, 213

Potential, thermodynamic, 59, 64  
 Pressure, 9, 55, 75, 106, 179, 193, 230, 235  
 atmospheric, 259, 268  
 partial, 73, 80  
 radiation, 55, 225  
 vapor, 73  
 Probability, 87, 149  
 Process  
 adiabatic, 29, 34, 45, 230, 251, 259  
 electrochemical, 82  
 irreversible, 43  
 isothermal, 29, 33  
 quasistatic, 24, 27  
 reversible, 34, 40  
 spontaneous, 42  
 Protons, 213, 243

Quantum states, 144, 167  
 statistics, 208, 239  
 Quasistatic process, 24, 27

Radiation, thermal, 54, 221, 265  
 fluctuation in, 226, 265  
 Random walk, 90, 129  
 Rayleigh-Jeans distribution, 223  
 Reactions, chemical, 78  
 Relaxation time, 116  
 Reservoir  
 heat, 26, 29, 34  
 work, 27  
 Reversible process, 34, 40  
 Rotational partition function, 195

Sakur-Tetrode formula, 189  
 Scale  
 centigrade, 7, 16  
 factor, 108  
 Kelvin, 17, 38  
 thermodynamic, 38, 40  
 Schrödinger's equation, 144, 183, 238, 266  
 Solid state, 19, 68, 111, 154, 170, 266

Space  
  momentum, 96  
  phase, 103, 141, 144  
Specific heat, 15, 172, 197, 201, 229, 236, 247  
Spin, 227, 233  
Spontaneous process, 42  
Standard deviation, 90, 98  
State  
  equation of, 16, 18, 64, 165, 179, 193, 204, 250  
  macro, 141  
  micro, 5, 141, 146  
  system, 185, 208  
  variables, 6, 13  
Stefan's law, 55, 225  
Stirling's formula, 95, 156  
Stokes' law, 131  
Stoichiometric coefficient, 78  
Sublimation, 74, 264  
Susceptibility, magnetic, 15  
Symmetric wave functions, 240, 267  
System state, 185, 208  
  
Taylor's series, 105  
Temperature, 7  
  Debye, 178  
  of melting, 70  
  scale, 7, 17, 38, 40  
  thermodynamic, 38, 40  
  of vaporization, 72  
Time, mean free, 102, 257  
Thermionic emission, 257, 266  
Translational partition function, 183  
Transport phenomena, 116  
Triple point, 74, 256  
  
Ultraviolet catastrophe, 223  
Uncertainty principle, 145  
Universe, entropy of, 44, 255  
  
Van der Waals' equation, 18, 48, 198, 251, 255  
Van't Hoff's equation, 82, 256  
Vapor pressure, 73  
Vaporization, latent heat of, 72  
Variables  
  dependent, 6  
  extensive, 14, 41, 59  
  independent, 6  
  intensive, 14, 41  
  state, 6, 13  
Variance, 90, 99  
Velocity distribution, 96, 160, 257  
  drift, 98, 119, 260  
  mean, 100  
  mean square, 123  
Vibrational partition function, 199  
  specific heat, 179, 200  
Vibrations  
  of crystal, 112, 155, 175, 200  
  of molecule, 199  
Virial equation, 18  
Viscosity, 131, 259  
  
Walk, random, 90, 129  
Wave functions, 183, 238  
Weight, statistical, 192  
Work, 23  
Work reservoir, 27



3 1262 04065 7776

536.7  
M 667  
-2

PHYSICS  
SEMINAR

ENGINEERING  
& PHYSICS  
LIBRARY

# A Series of Lecture Note and Reprint Volumes

**NUCLEAR MAGNETIC RELAXATION: A Reprint  
Volume**

by N. Bloembergen      192 pages      \$3.95

**S-MATRIX THEORY OF STRONG INTERACTIONS:**

A Lecture Note and Reprint Volume

by Geoffrey F. Chew      192 pages      \$3.95

**QUANTUM ELECTRODYNAMICS: A Lecture Note**

and Reprint Volume

by R. P. Feynman      208 pages      \$3.95

**THE THEORY OF FUNDAMENTAL PROCESSES:**

A Lecture Note Volume

by R. P. Feynman      160 pages      \$3.95

**THE MOSSBAUER EFFECT: A Collection of Re-  
prints with an Introduction**

by Hans Frauenfelder      256 pages      \$3.95

**THE MANY-BODY PROBLEM: A Lecture Note and  
Reprint Volume**

by David Pines      464 pages      \$3.95

**QUANTUM THEORY OF MANY-PARTICLE SYS-  
TEMS: A Lecture Note and Reprint Volume**

by L. Van Hove, N. M. Hugenholtz,  
and L. P. Howland      272 pages      \$3.95

Order from your bookseller or write to:

**W. A. BENJAMIN, INC., PUBLISHERS**

2465 Broadway, New York 25, N. Y.